

Aquatic Informatics, Inc

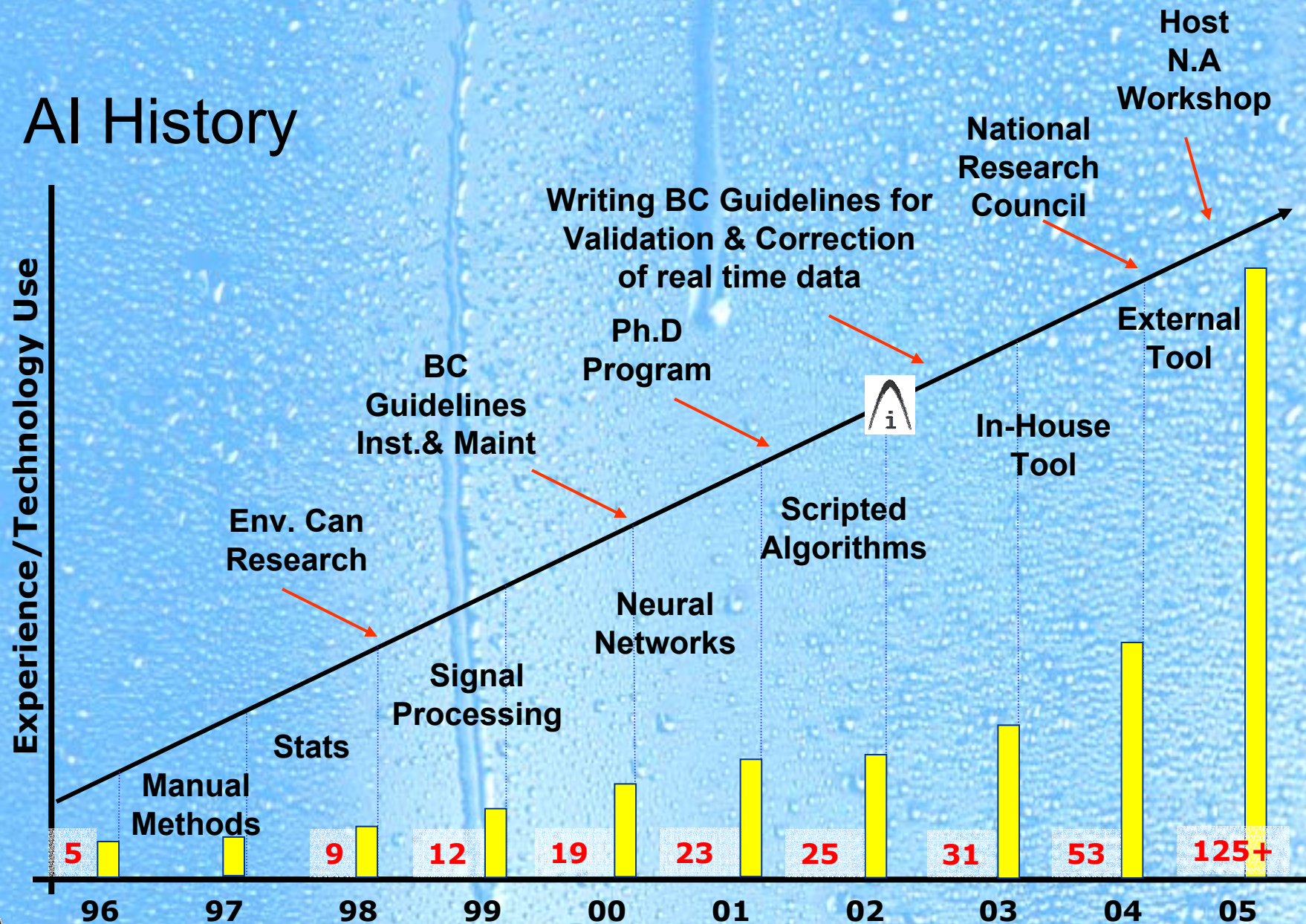
Data Validation and Correction Tools for the Real Time Water Quality Monitoring Industry

November 8th, 2004

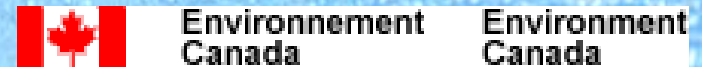
Portland, Oregon



AI History



AI Customers



AI Scientific Advisory Board

- Paul Whitfield –Environment Canada
- Dr. Dan Moore – UBC Hyrdology
- Dr. William Hsieh – UBC Climate Group
- Dr. John Richardson – UBC Stream Ecology
- Dr. Sean Fleming – UBC Hydrogeology
- Dr. Harold Davis – UBC Engineering Physics
- Ed Quilty, Ph.D Cand. – Aquatic Informatics
- Andre Viljoen – Avteq Technology
- Joel Bellenson – Biocad



AI has the only tool today that cost-effectively manages data for the unique challenges of the real time water quality monitoring industry.



AI Tools

- Detect anomalous data.
- Improved efficiency
- Validated data for further analysis.
- Correct data using local knowledge.
- Real-time applications or warehoused data.

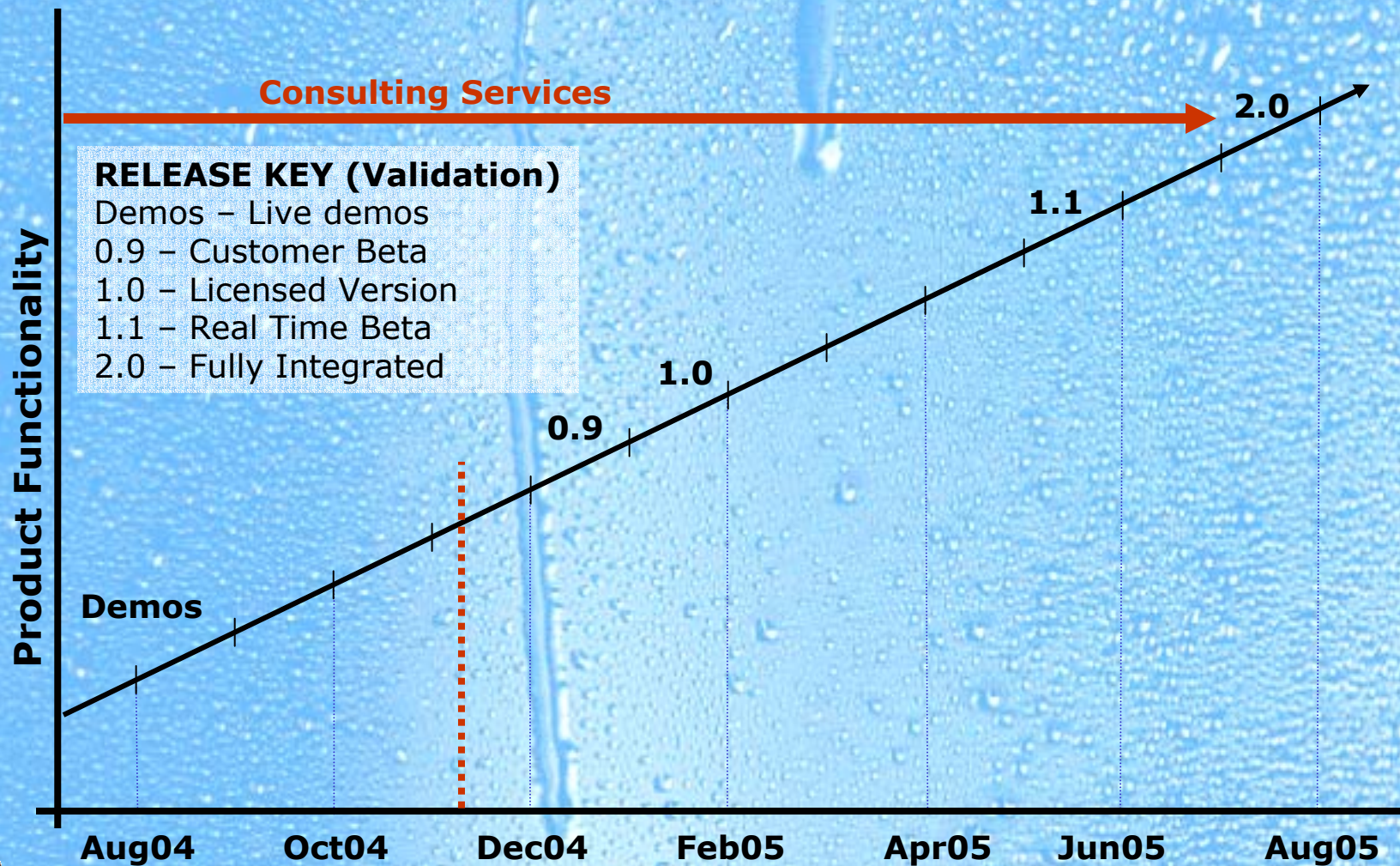


AI Customer Benefits

1. Reduction of errors
2. Confidence in data
3. Feedback into Monitoring Effectiveness
4. Better Information = Better Decisions
 - Improved Management of Industry
 - Public Safety
 - Improved Habitat
 - Proper Infrastructure



AI Roadmap



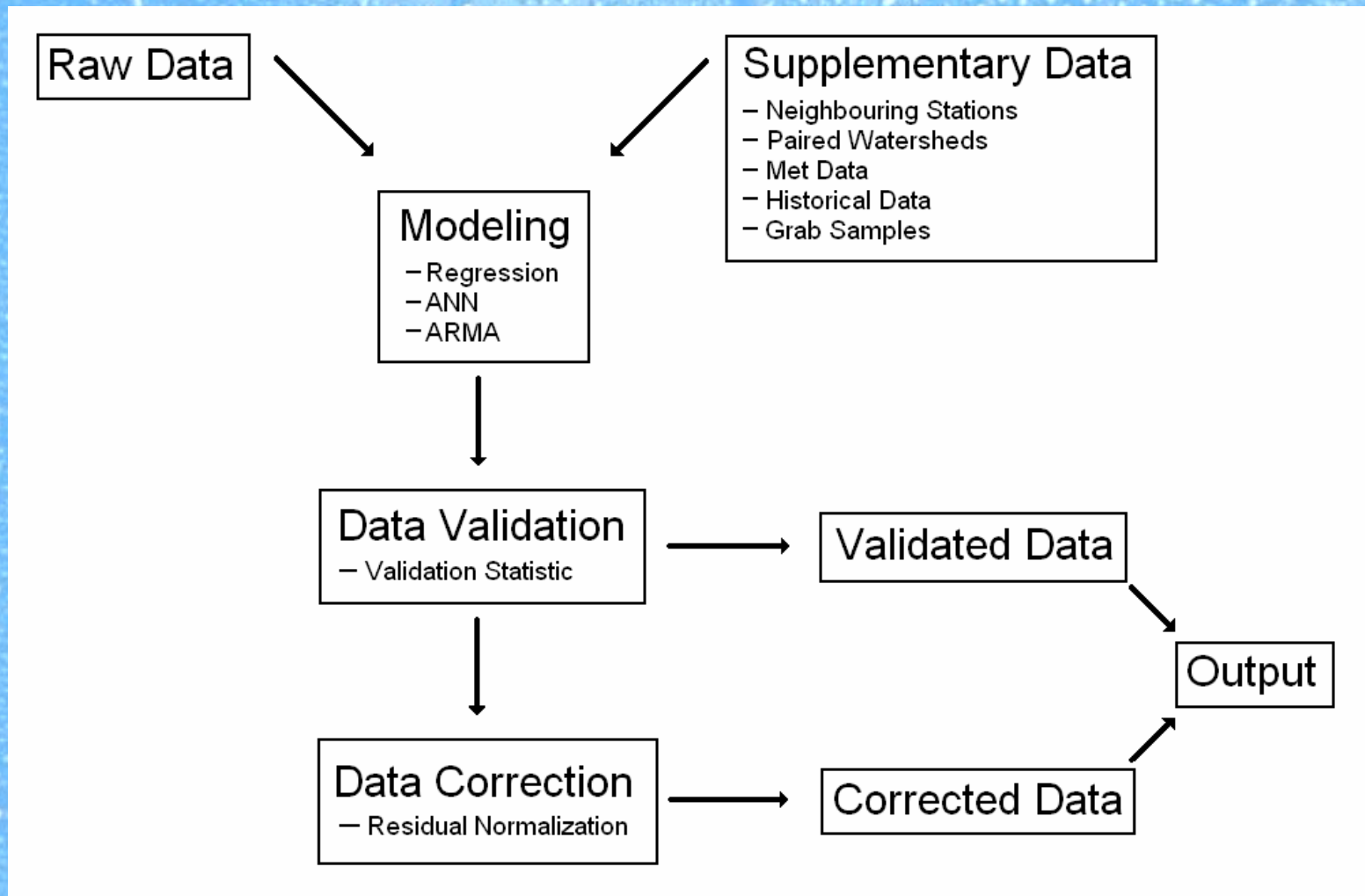
High Frequency Hydrologic Data Validation and Correction

Peter Hudson

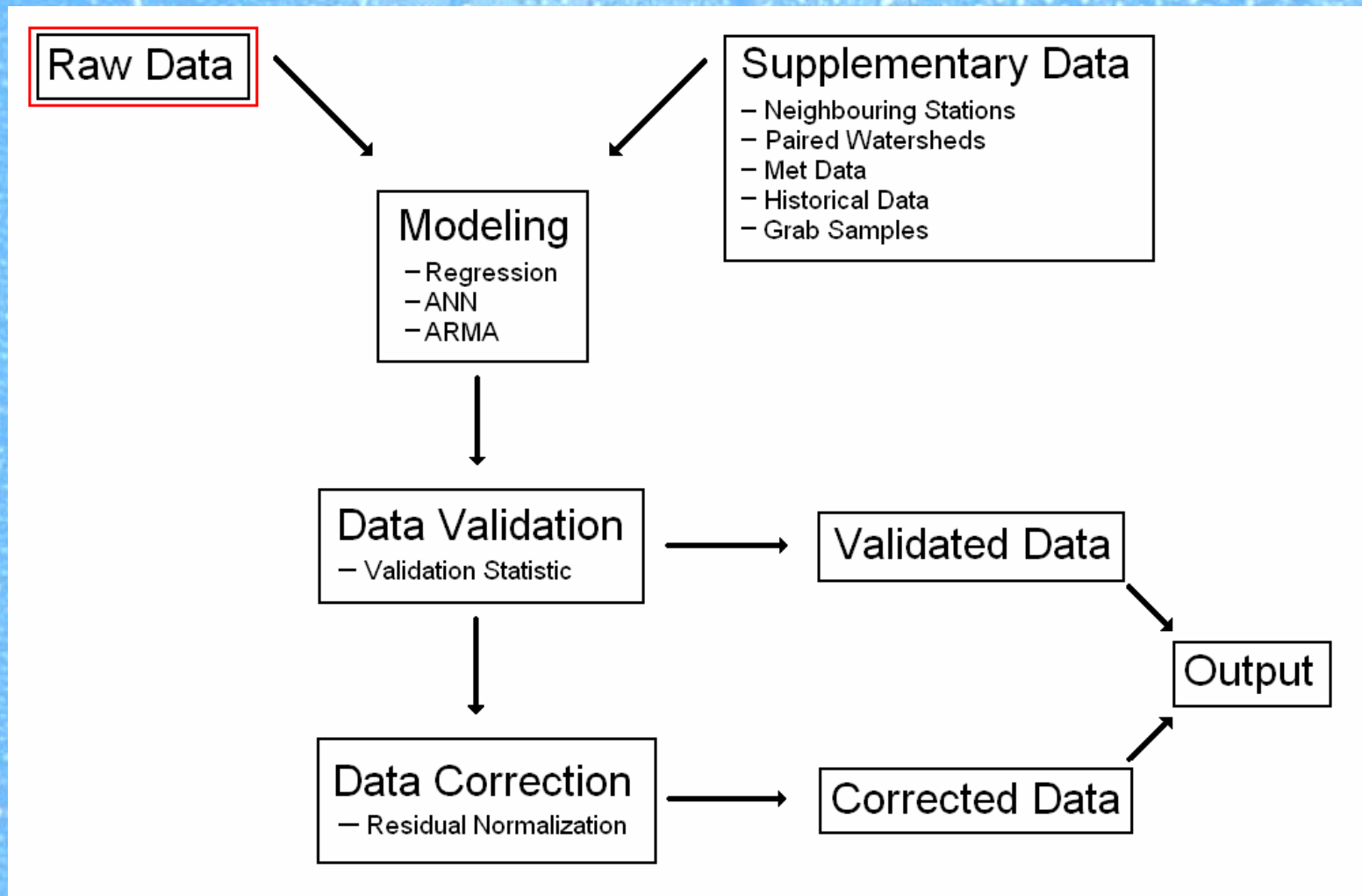
Senior Project Scientist



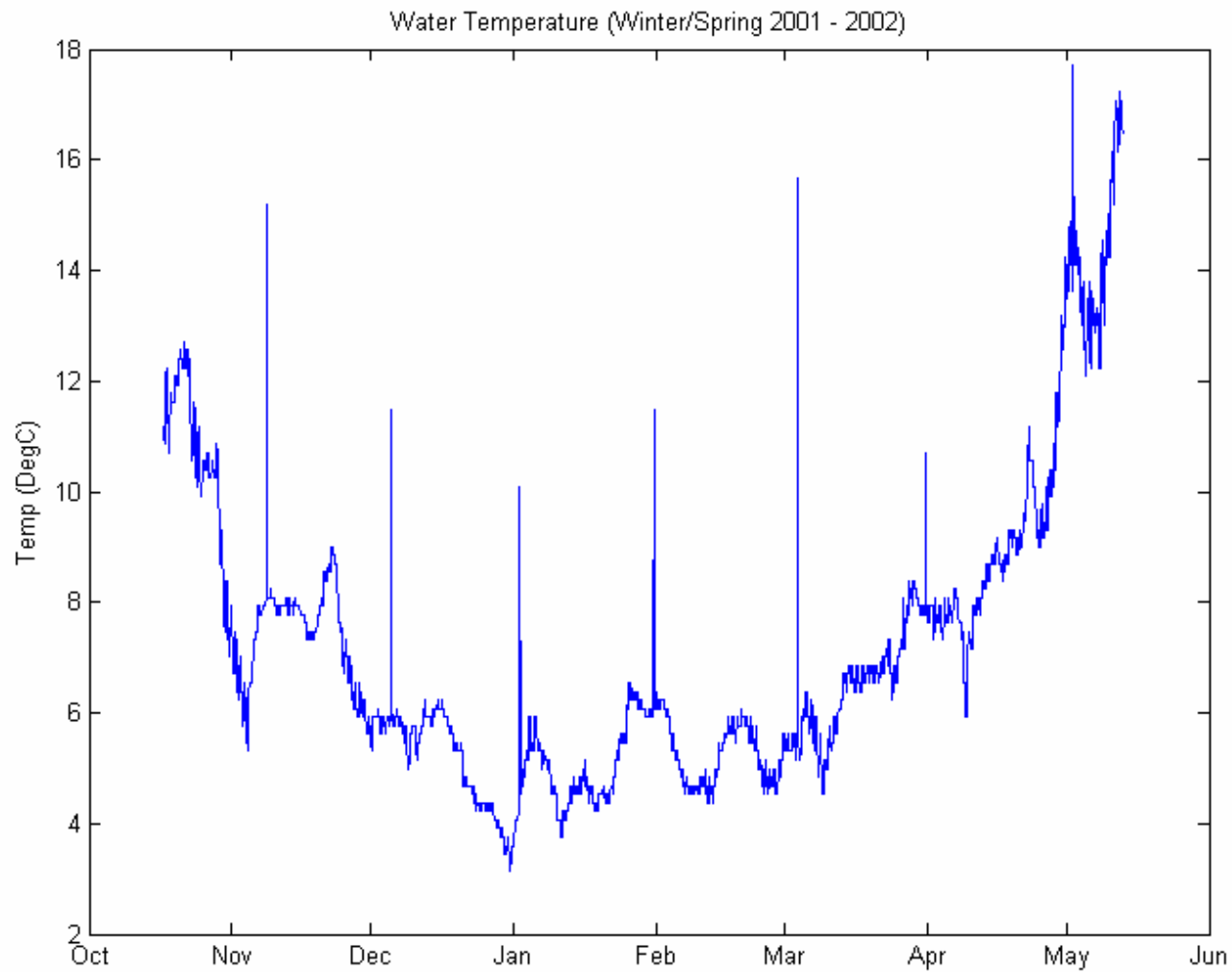
Data Management Flow Chart:



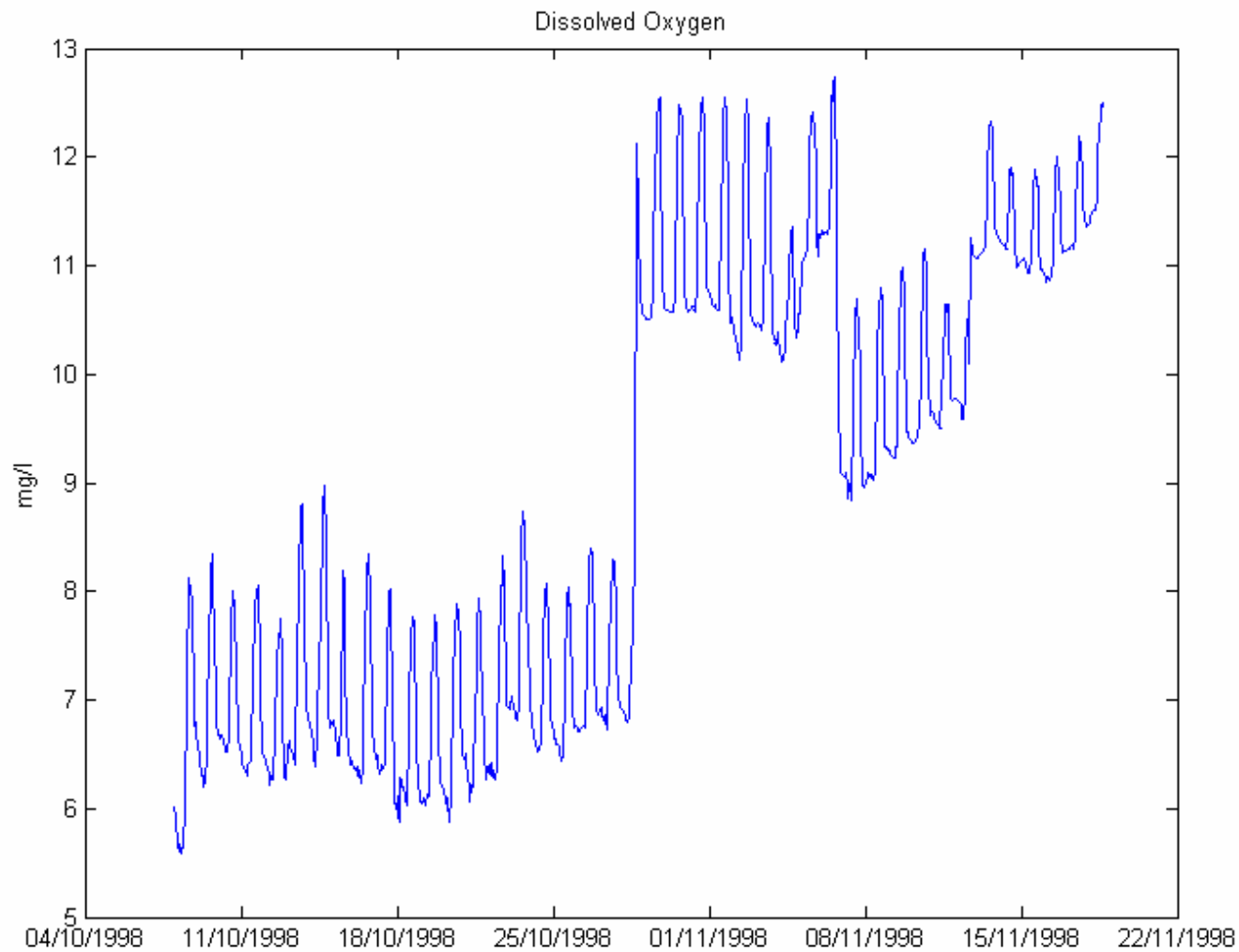
Data Management Flow Chart: Raw Data



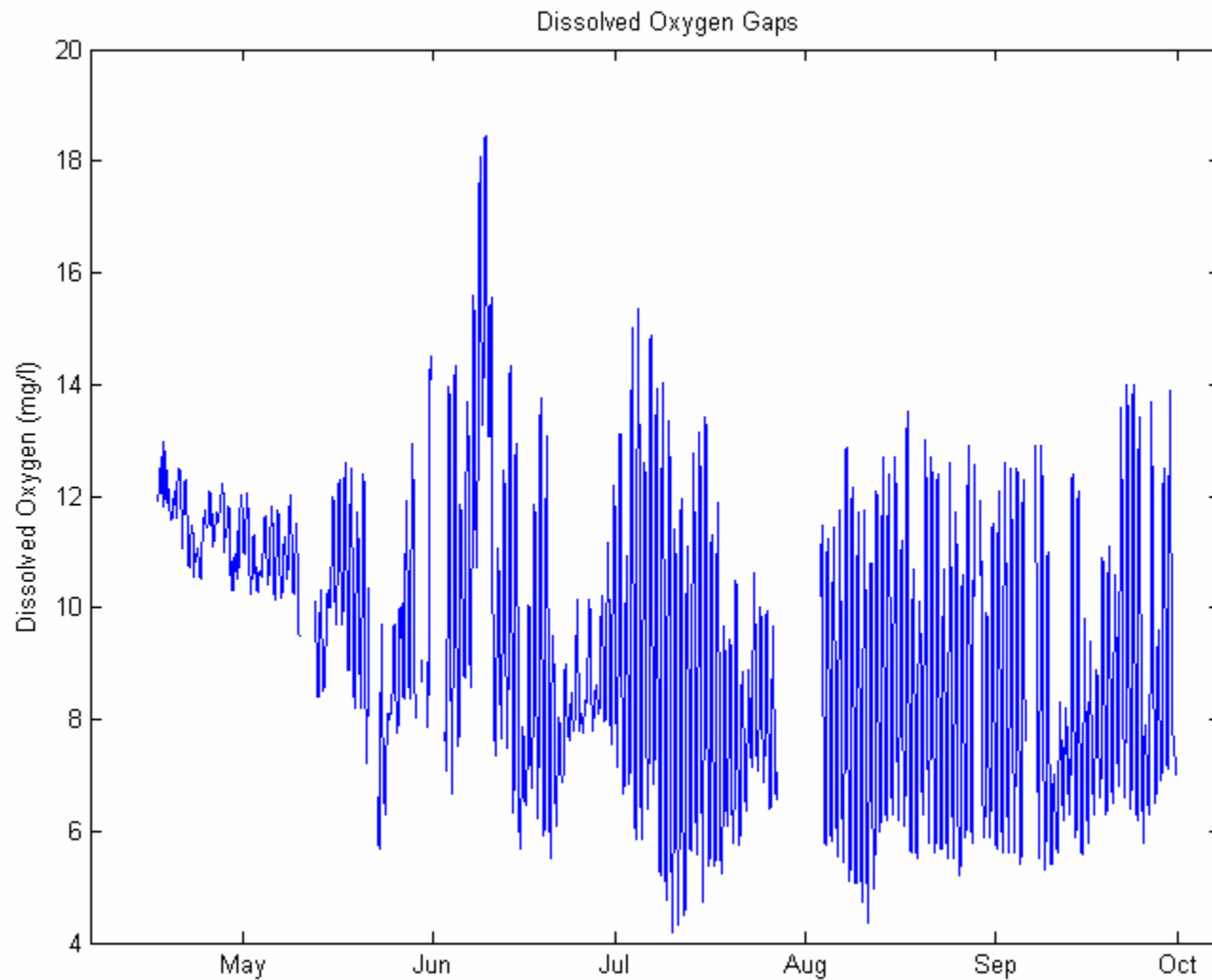
Raw Data: Outliers



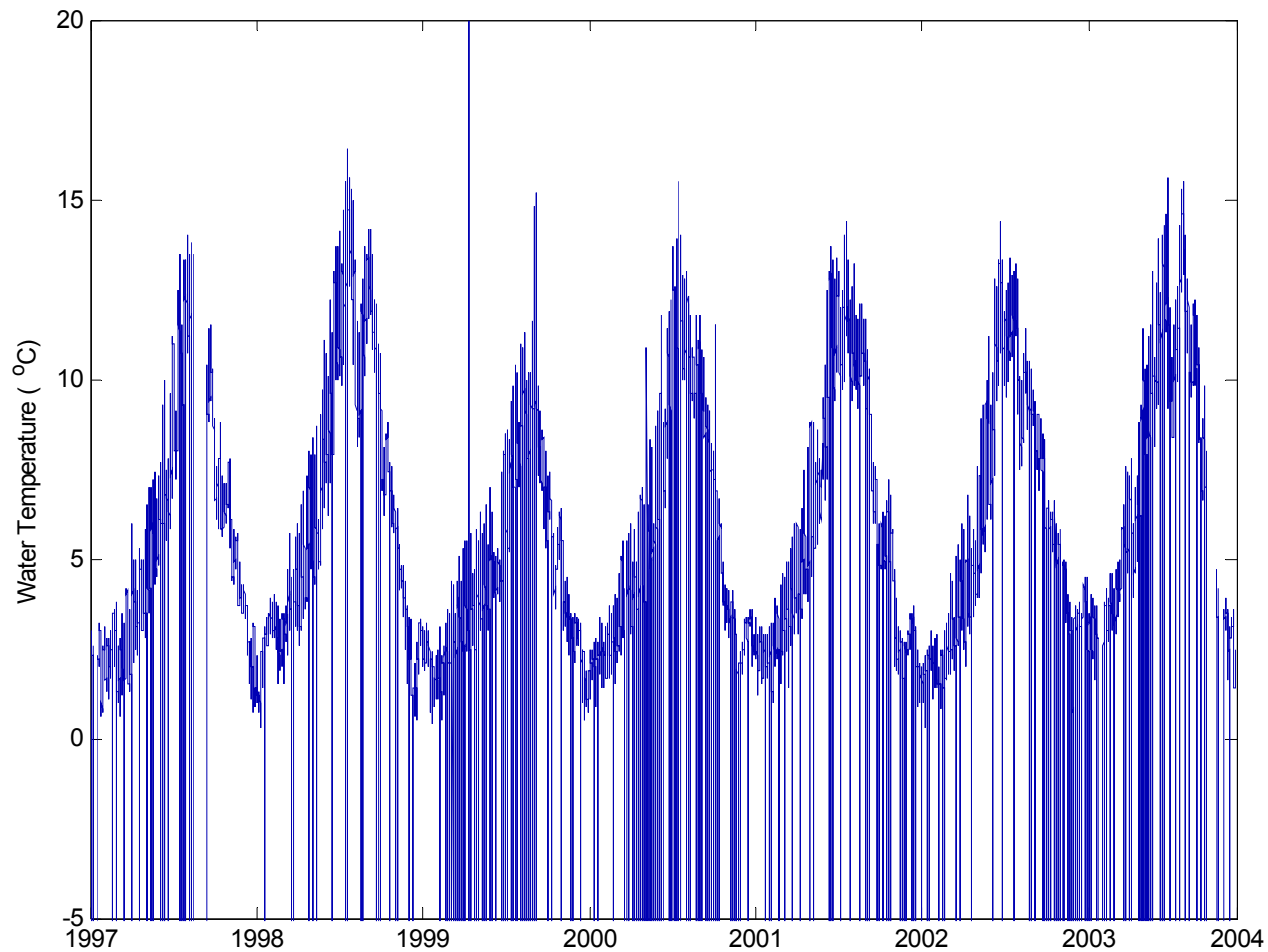
Raw Data: Calibration and Drift



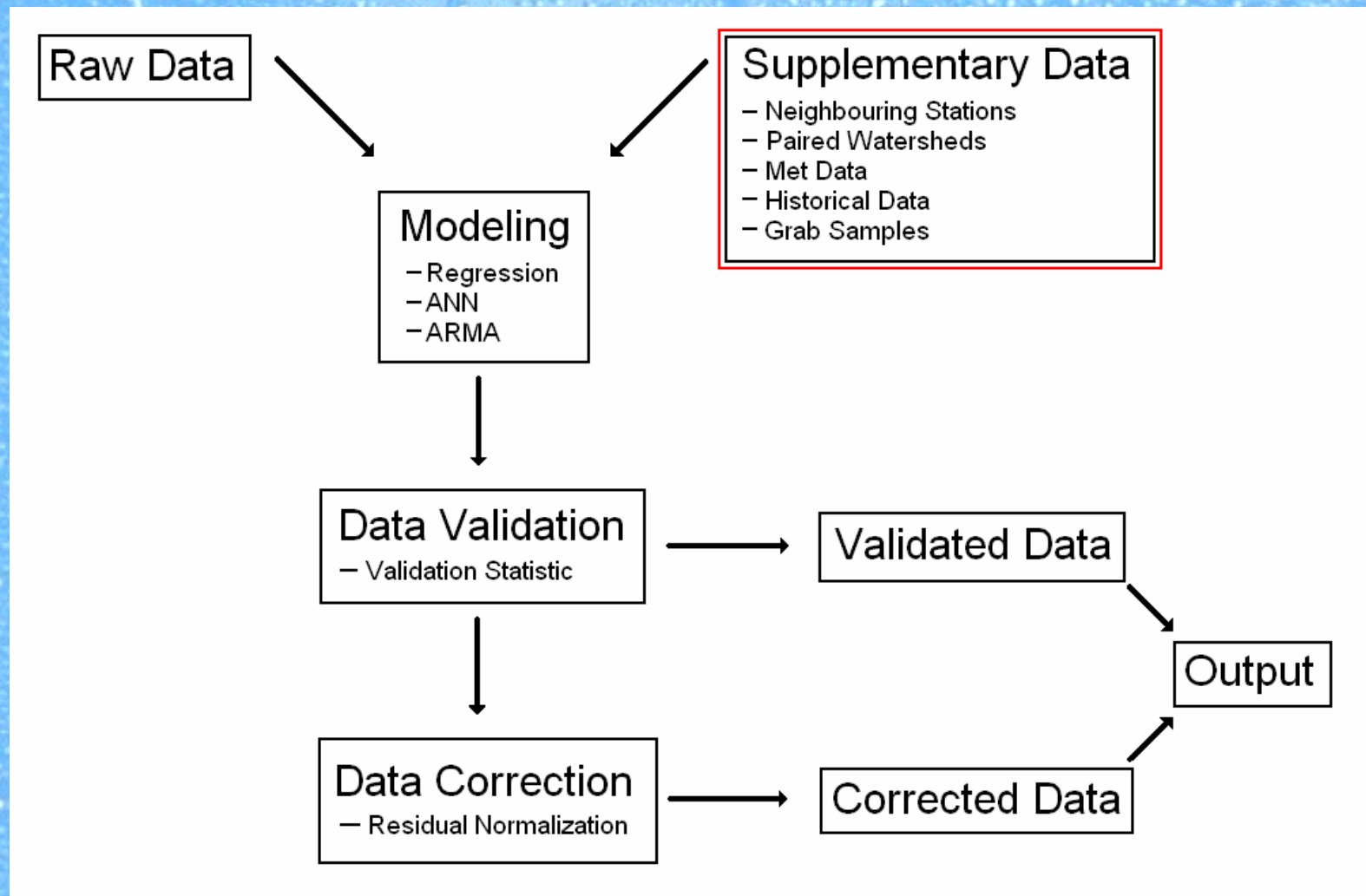
Raw Data: Gaps



Raw Data: Everything



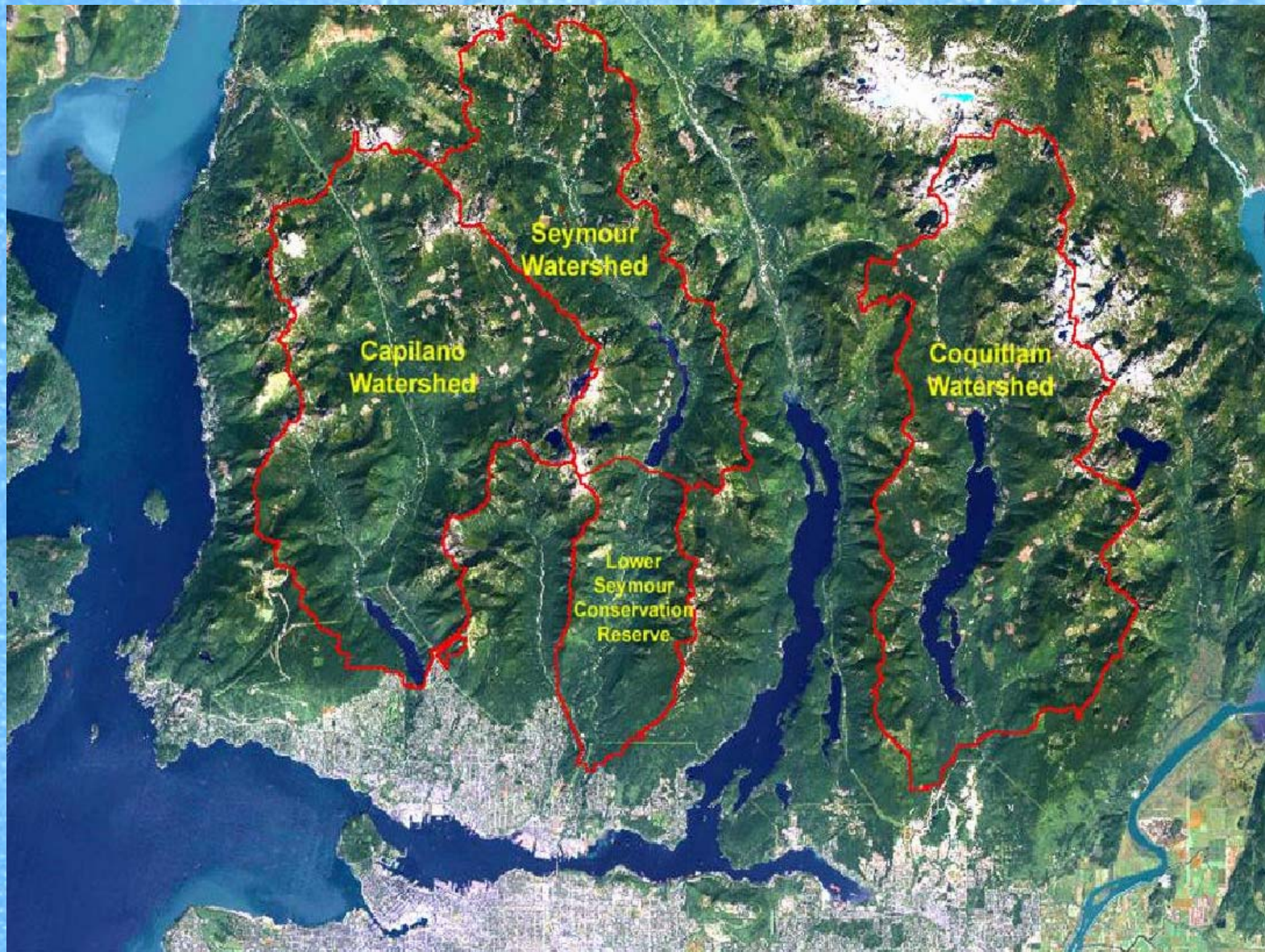
Data Management Flow Chart: Sup. Data



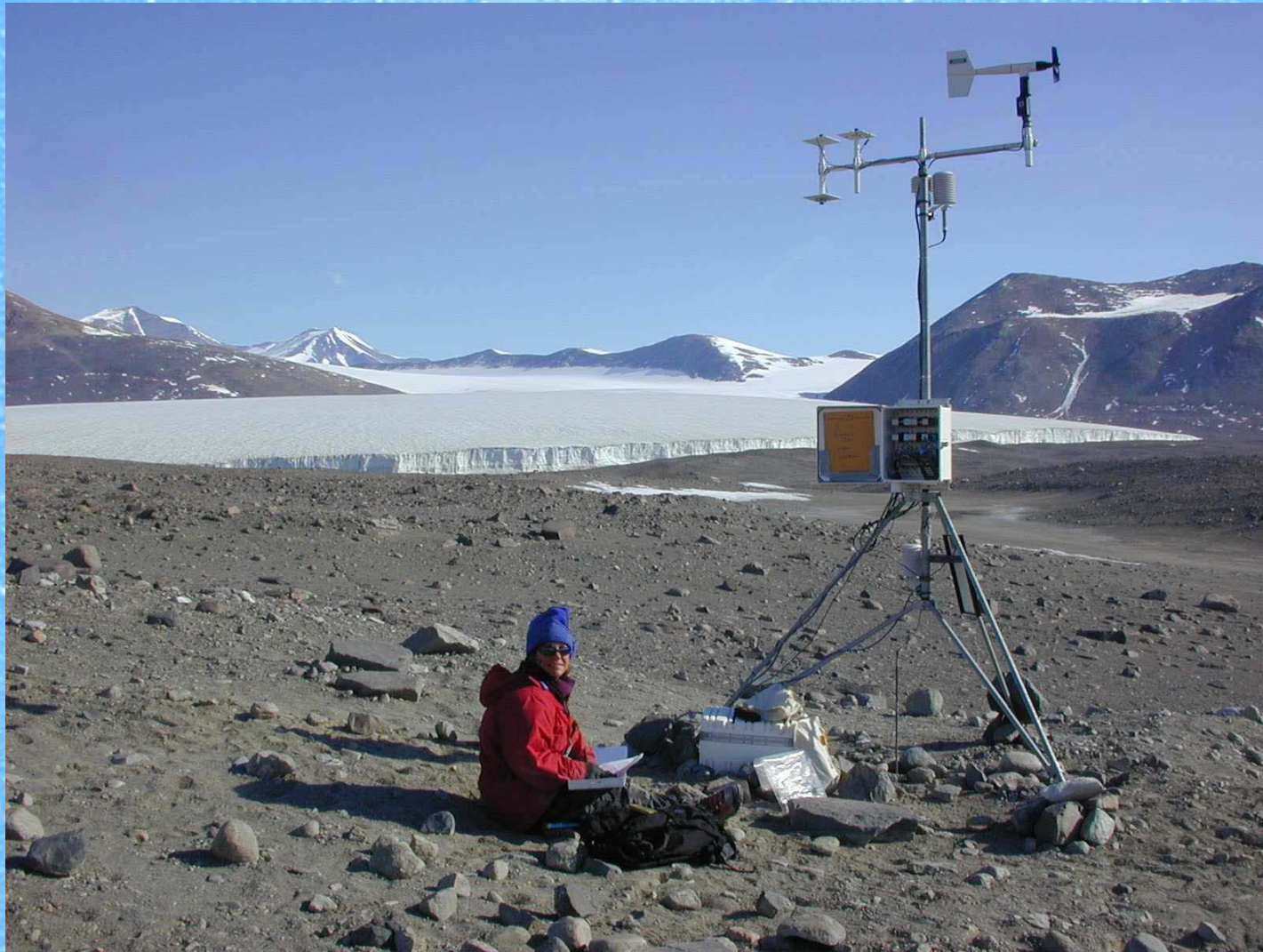
Supplementary Data: Neighbouring Stations



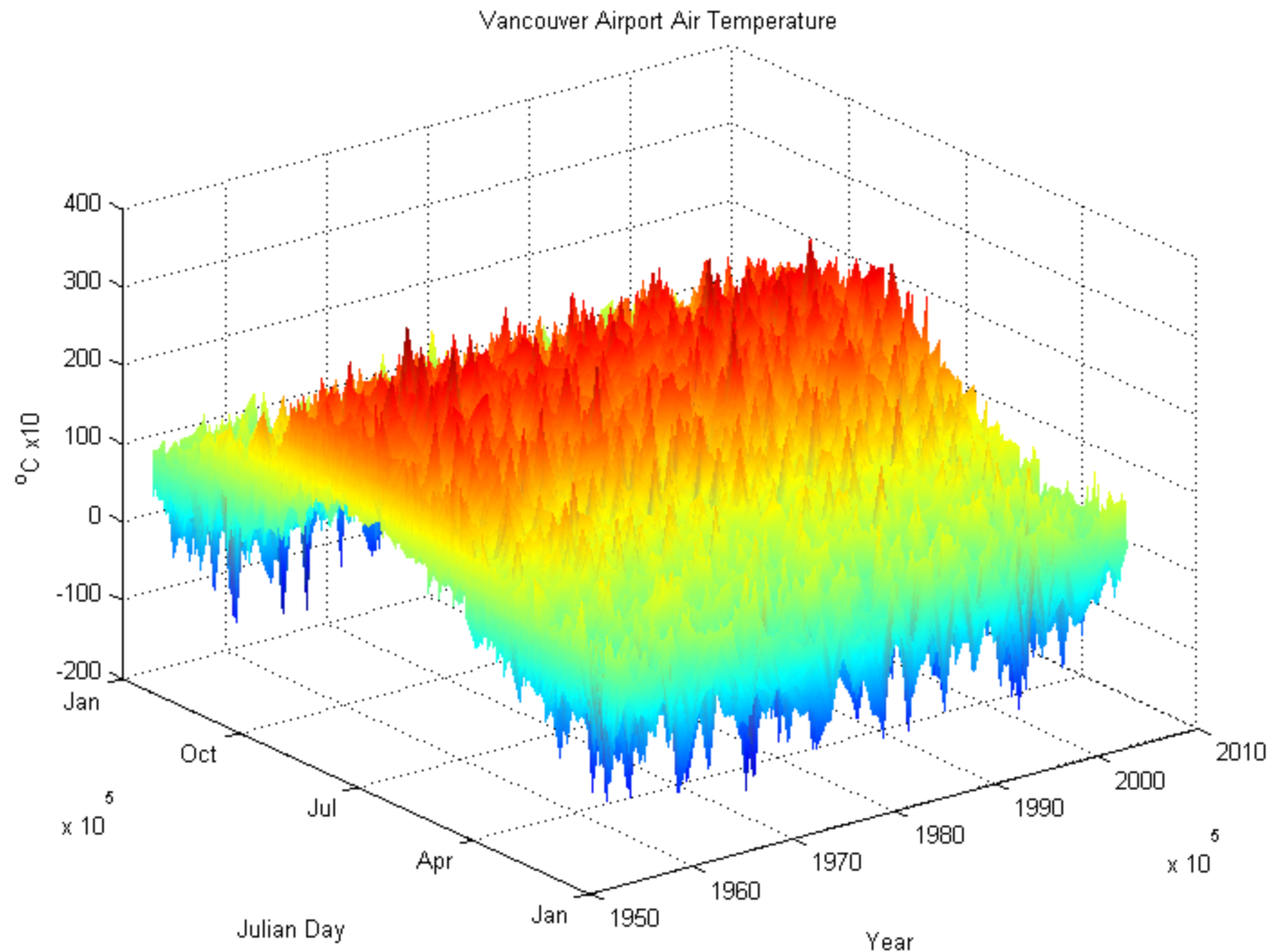
Supplementary Data: Paired Watersheds



Supplementary Data: Meteorological Data



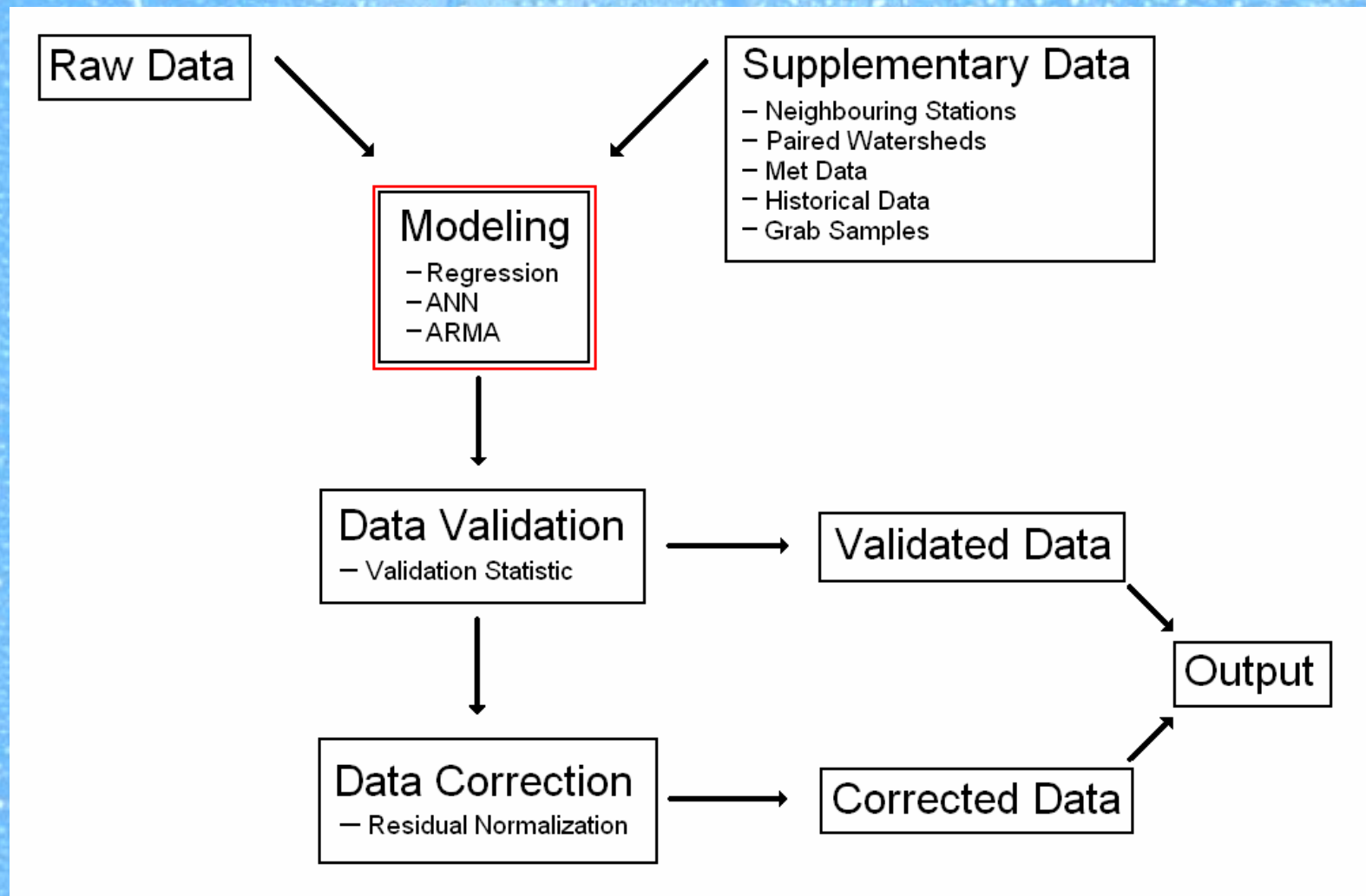
Supplementary Data: Historical Data



Supplementary Data: Grab Samples

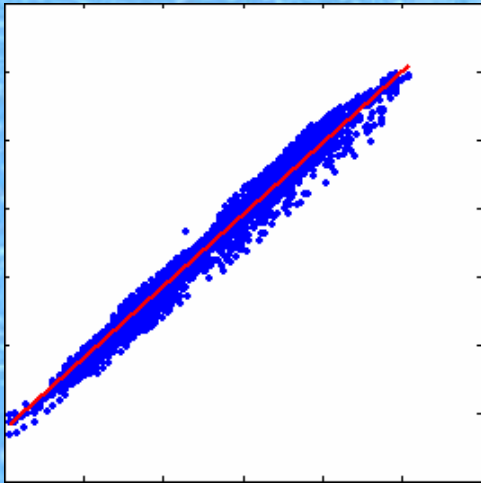


Data Management Flow Chart: Modeling

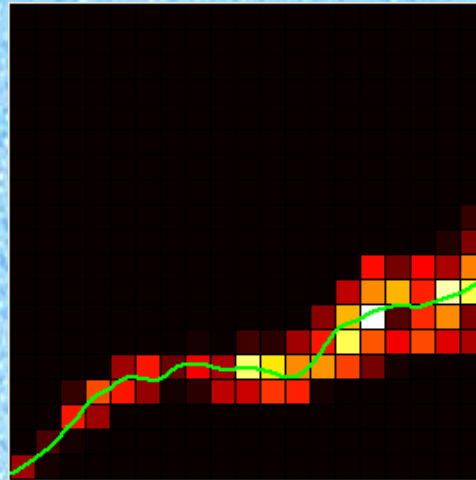


Modeling

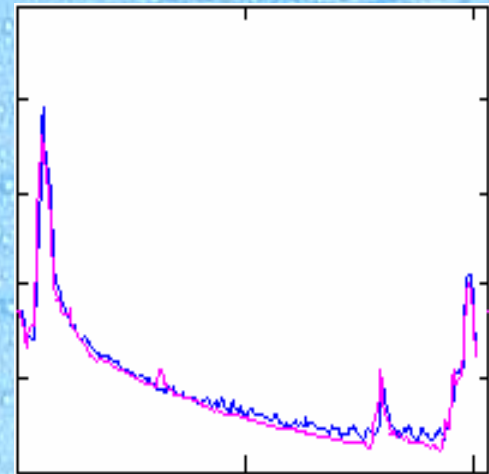
Multiple Regression &
Robust Regression



Artificial Neural
Networks

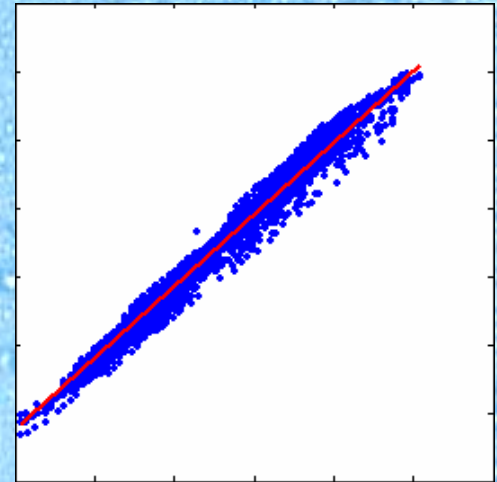


Auto-regressive
Moving average
Processes



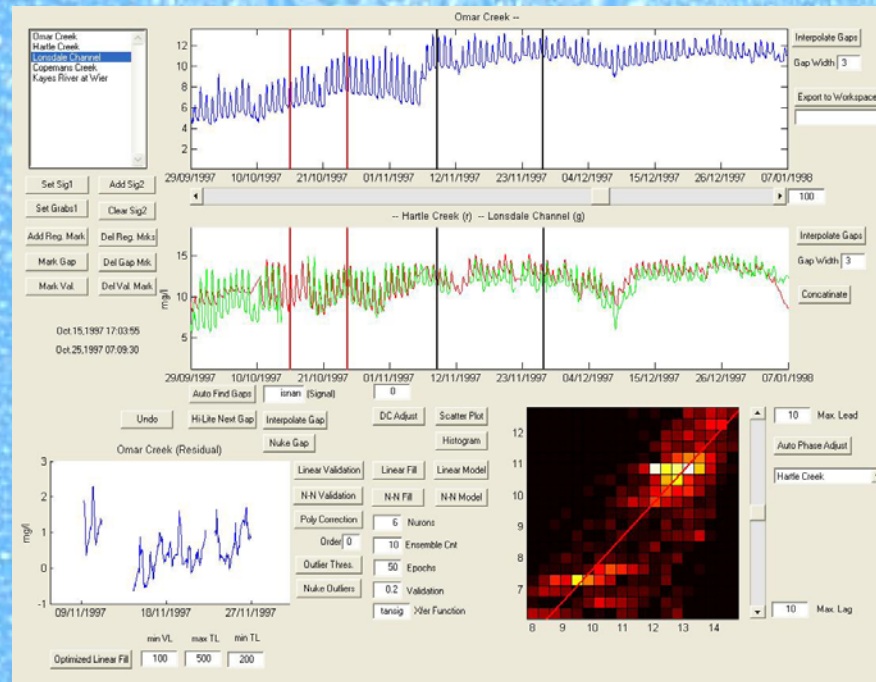
Modeling: Multiple Regression

- Coupled Watersheds / Multiple Stations
- Identify Areas of *good* data
- Maximize linear correlation by adjusting signal phase to account for physical factors (travel time, solar radiation, precipitation) and biological factors (photosynthesis and cellular decomposition).
- Build a multiple regression model

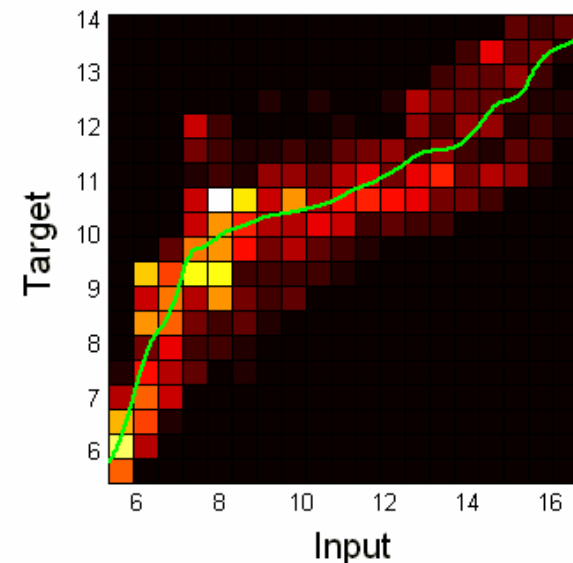
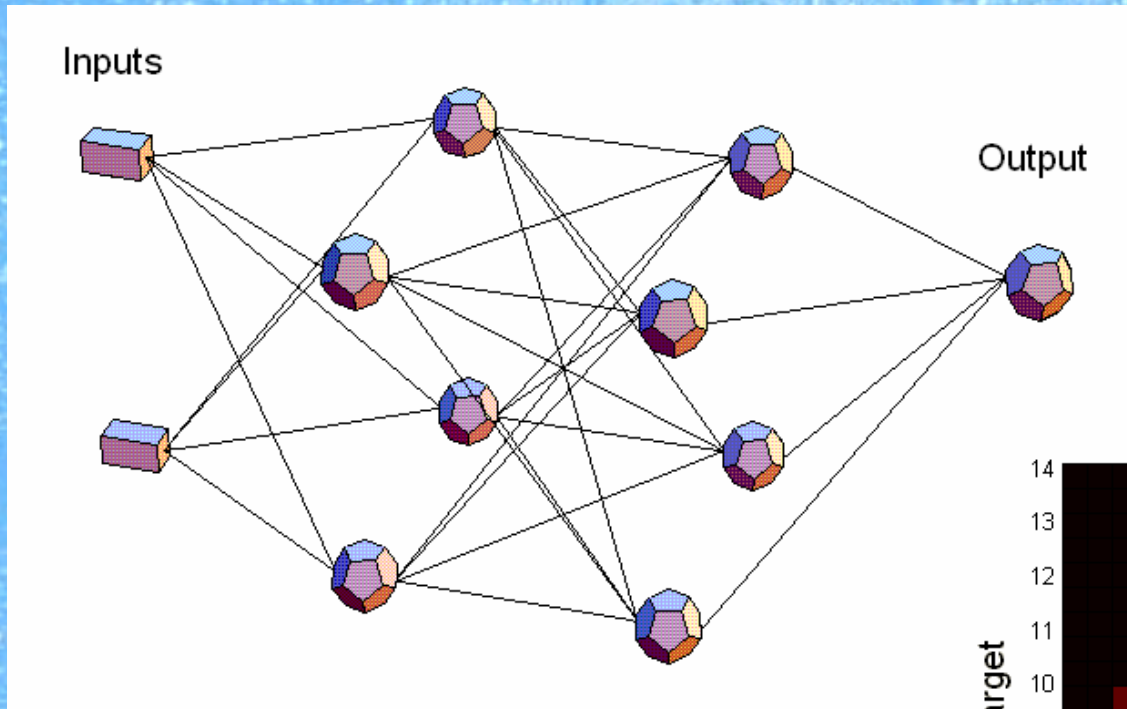


Modeling: Multiple Regression

Lets do it...

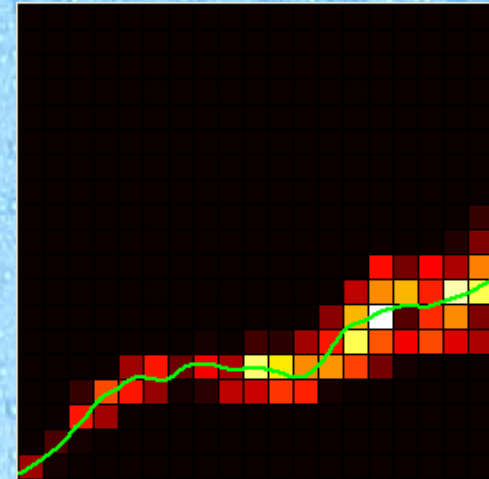


Modeling: Artificial Neural Networks



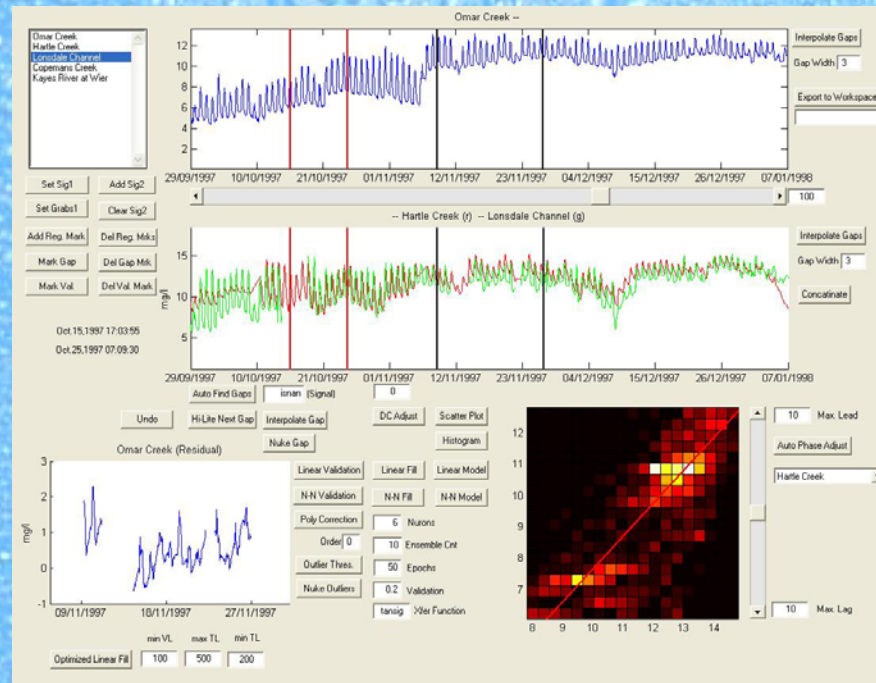
Modeling: Artificial Neural Networks

- Identify Areas of *good* data
- Maximize scatter density by adjusting signal phase to account for physical (travel time, solar radiation, precipitation) and biological (photosynthesis and cellular decomposition) factors.
- Build a Neural Network model

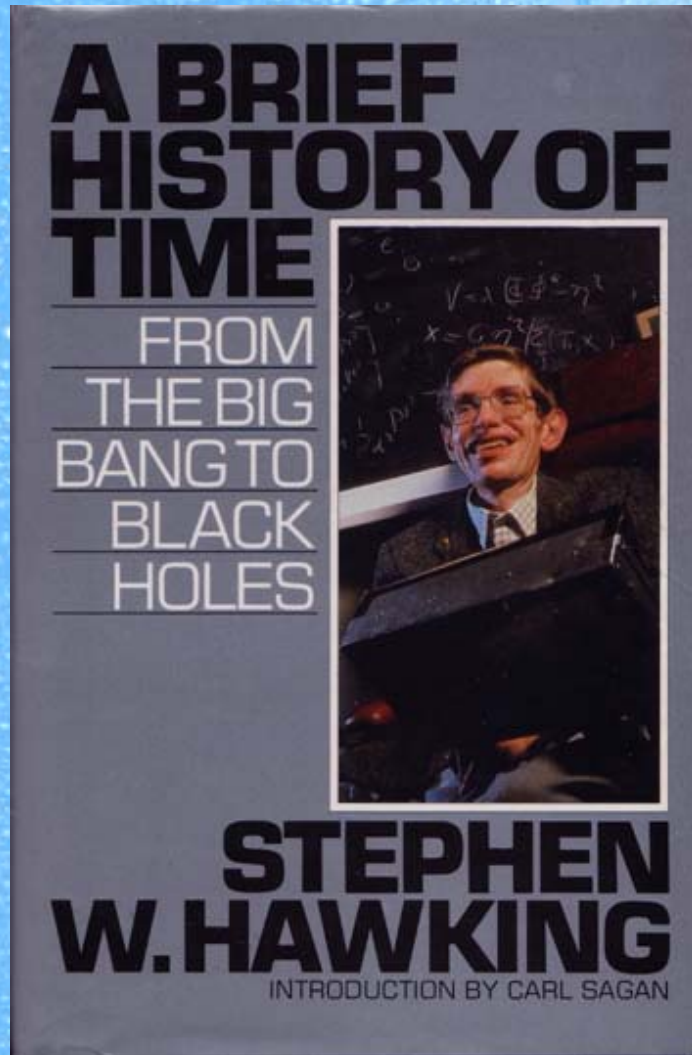


Modeling: Artificial Neural Networks

Let's do it...

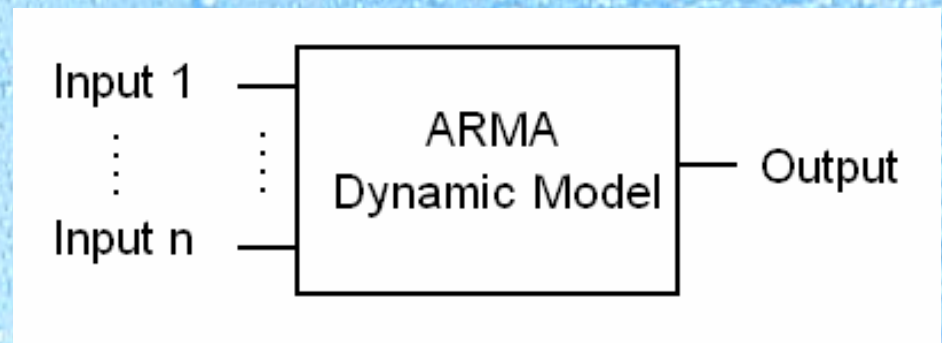


Modeling: Auto-Regressive Moving Average Processes

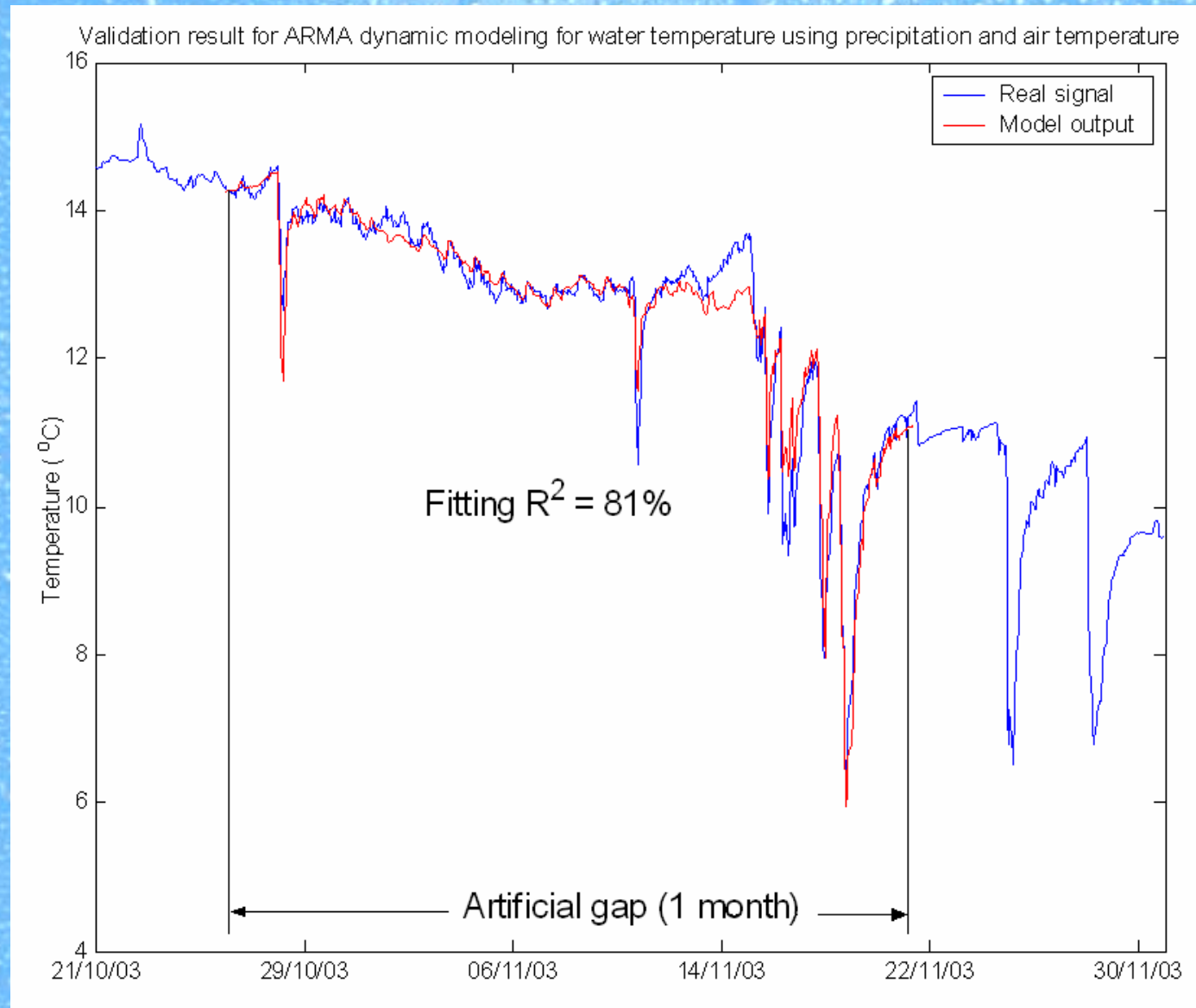


“... my publisher told me that for every equation I included I would half my book sales.”

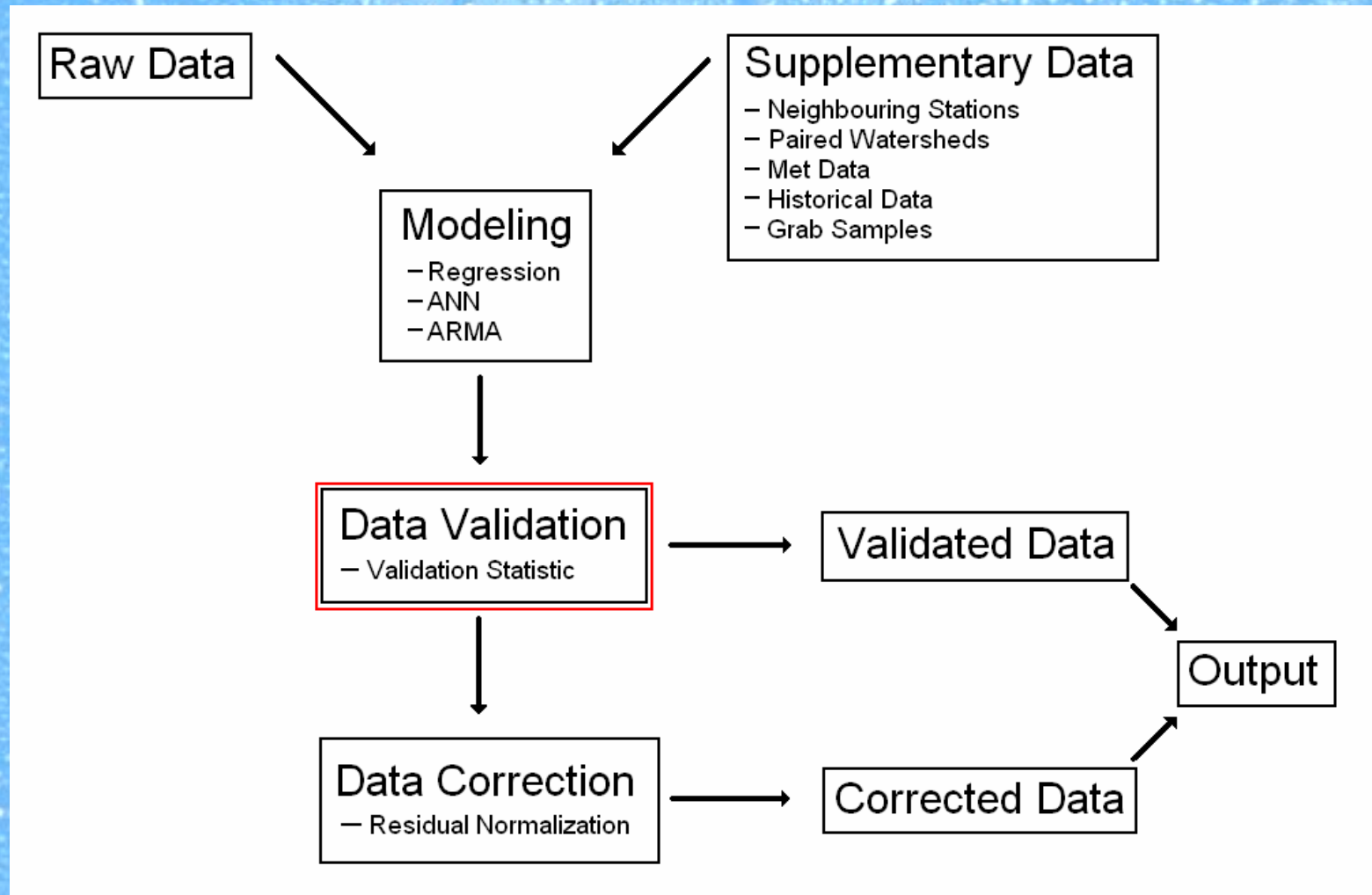
$$y_n = \sum_{k=1}^P a_k y_{n-k} + \sum_{m=0}^Q b_m x_{n-m} + \dots + c$$



Modeling: Auto-Regressive Moving Average Processes



Data Management Flow Chart: Data Validation

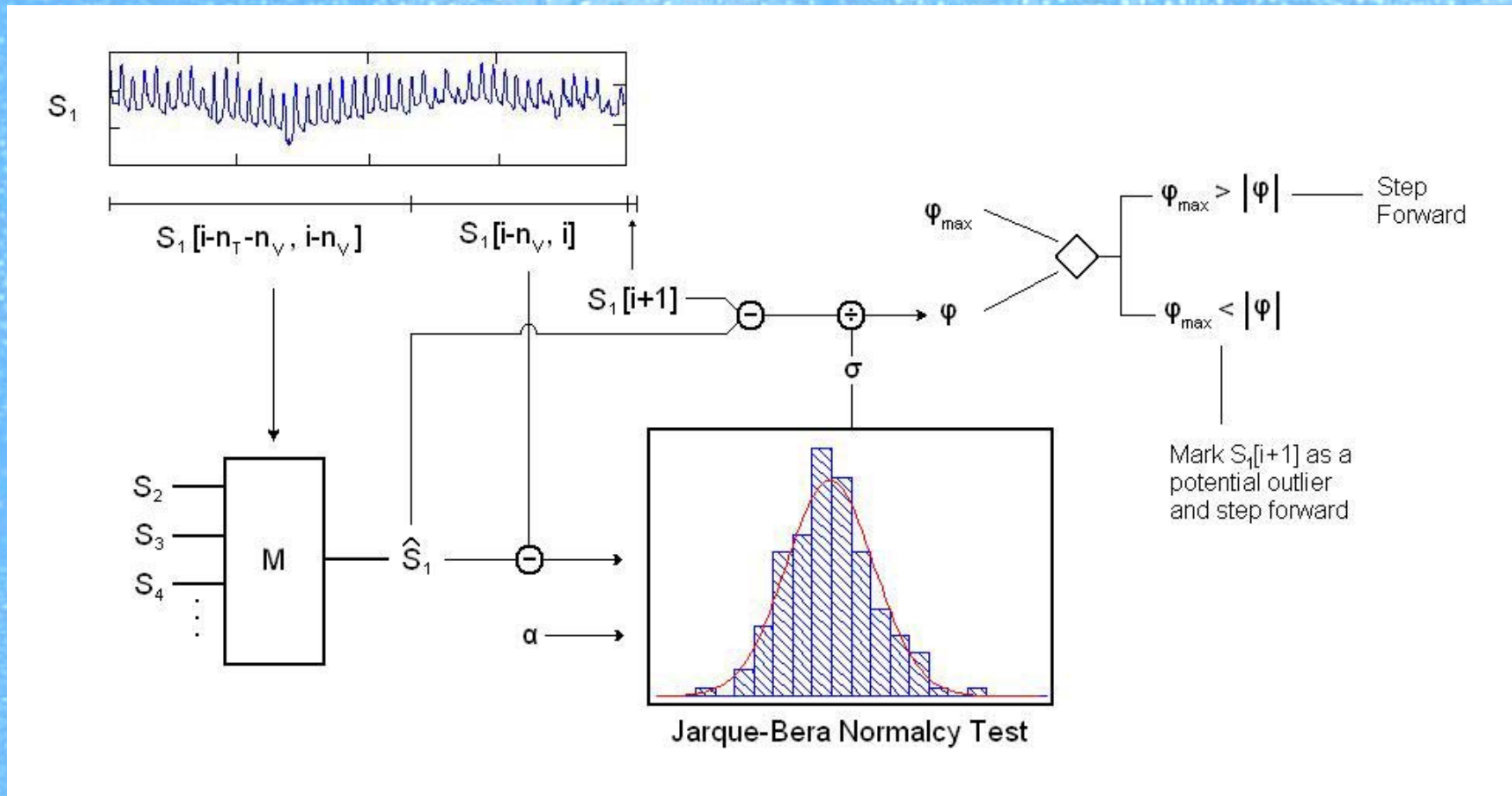


Data Validation: Validation Statistic (Conceptually)

- Identify some *good* 'warm up' data
- Calculate the expected value for the next data point
- Compare what the model predicted with what was actually measured
- Compute a validation statistic
- Step forward and repeat



Data Validation: Validation Statistic (Mathematically)

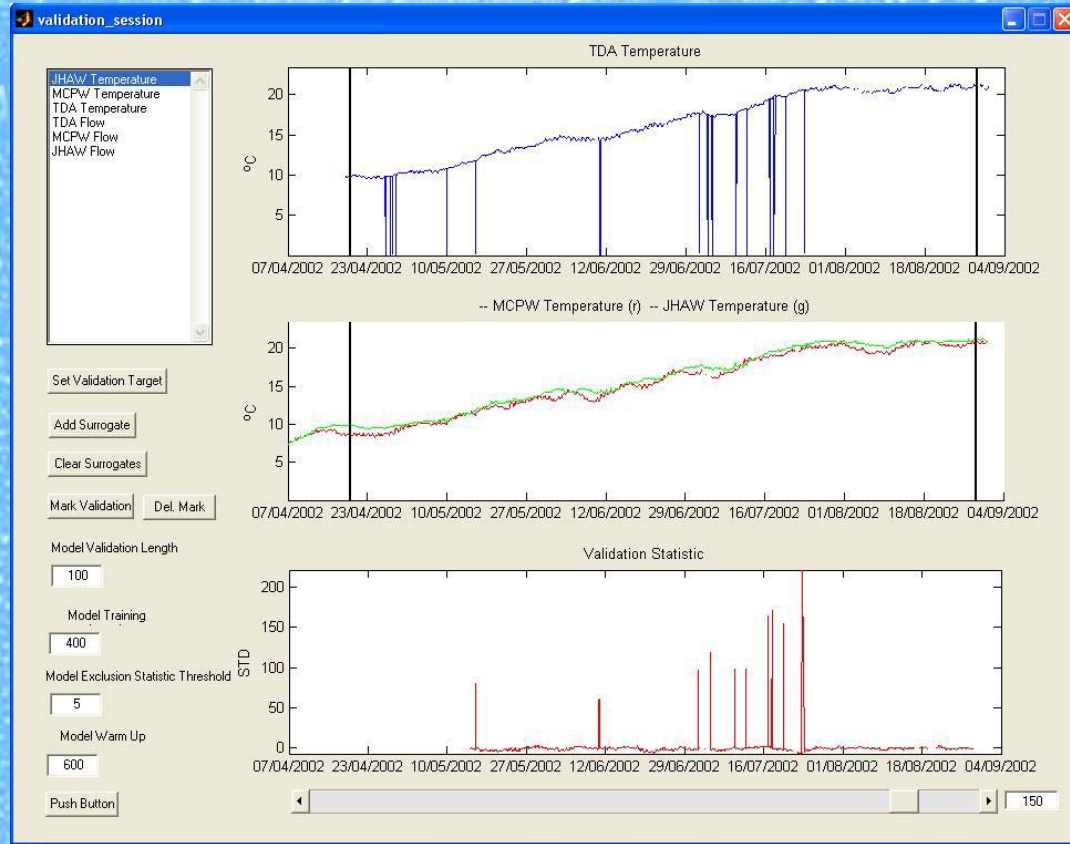


(See Validation Statistic White Paper)

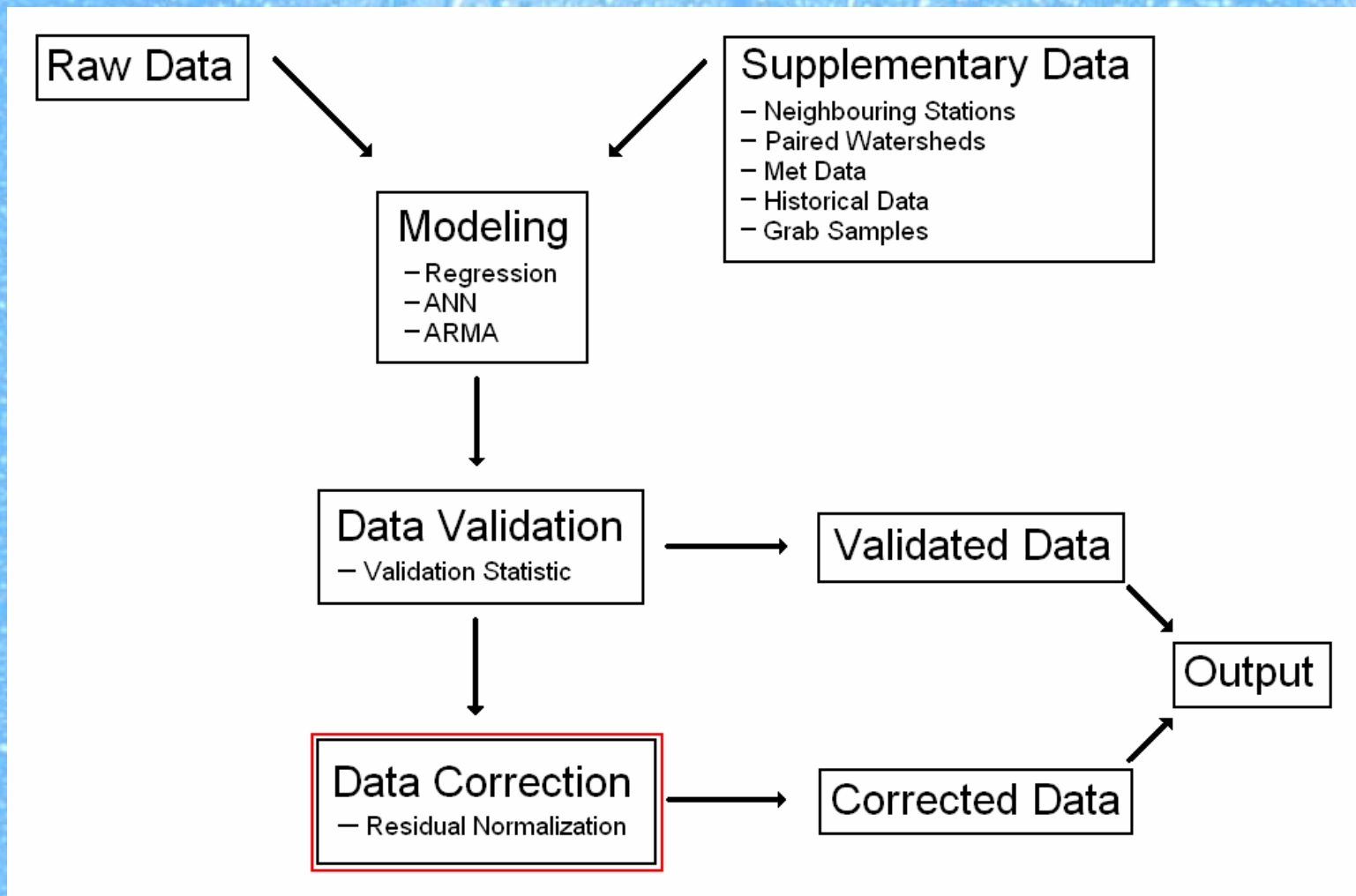


Data Validation: Validation Statistic

Let's Do it...

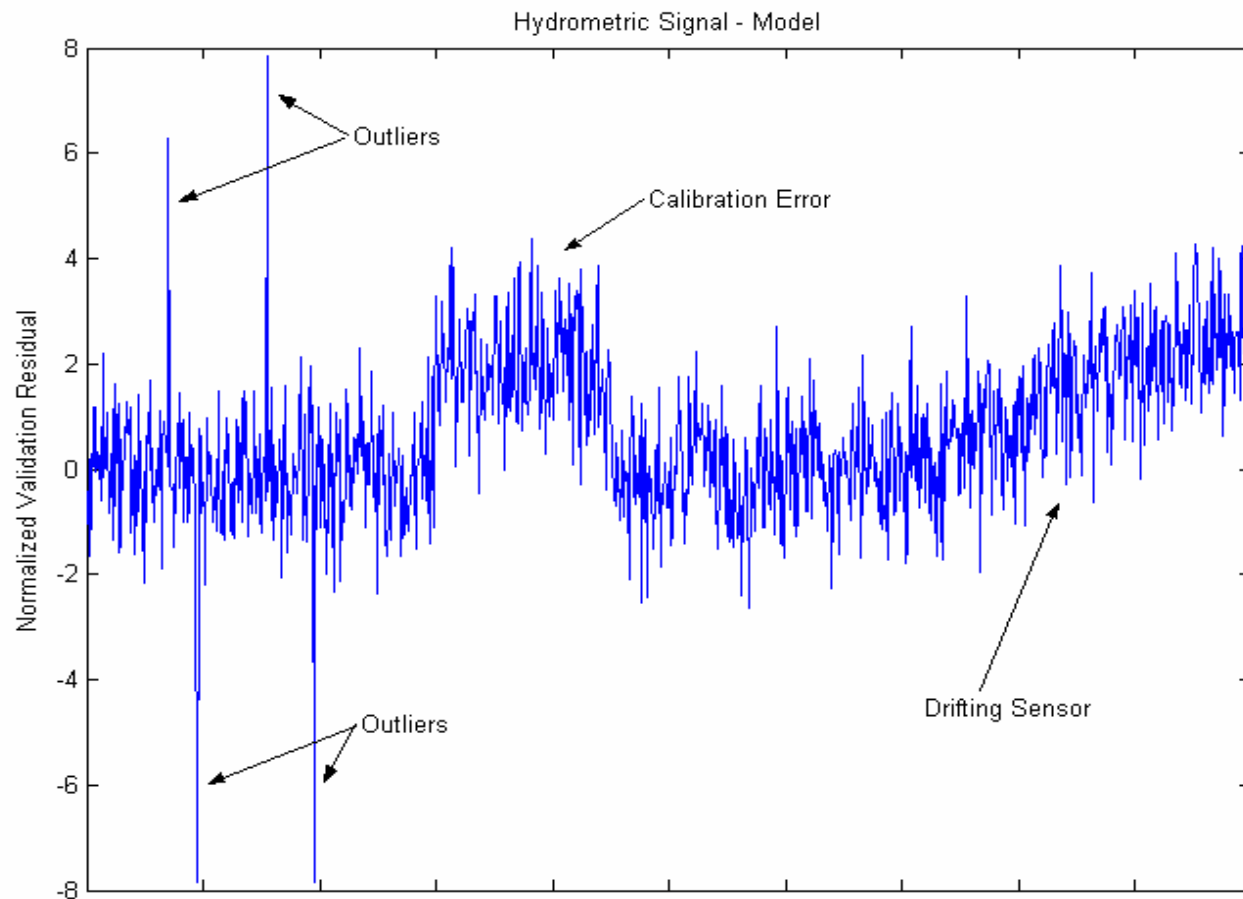


Data Management Flow Chart: Data Correction



Data Correction: Model Residual Normalization

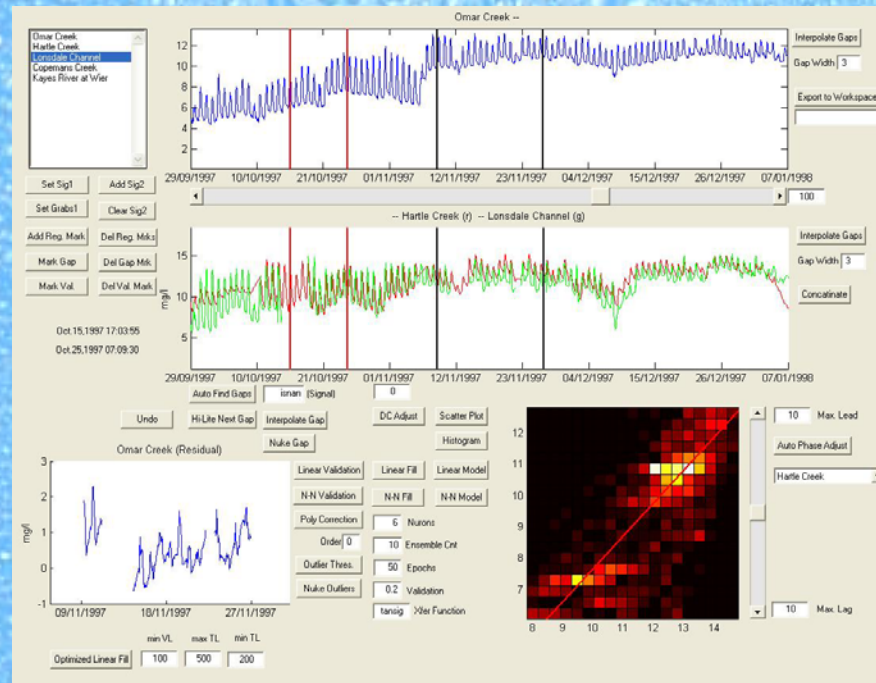
- Residual – White Noise Colouration:



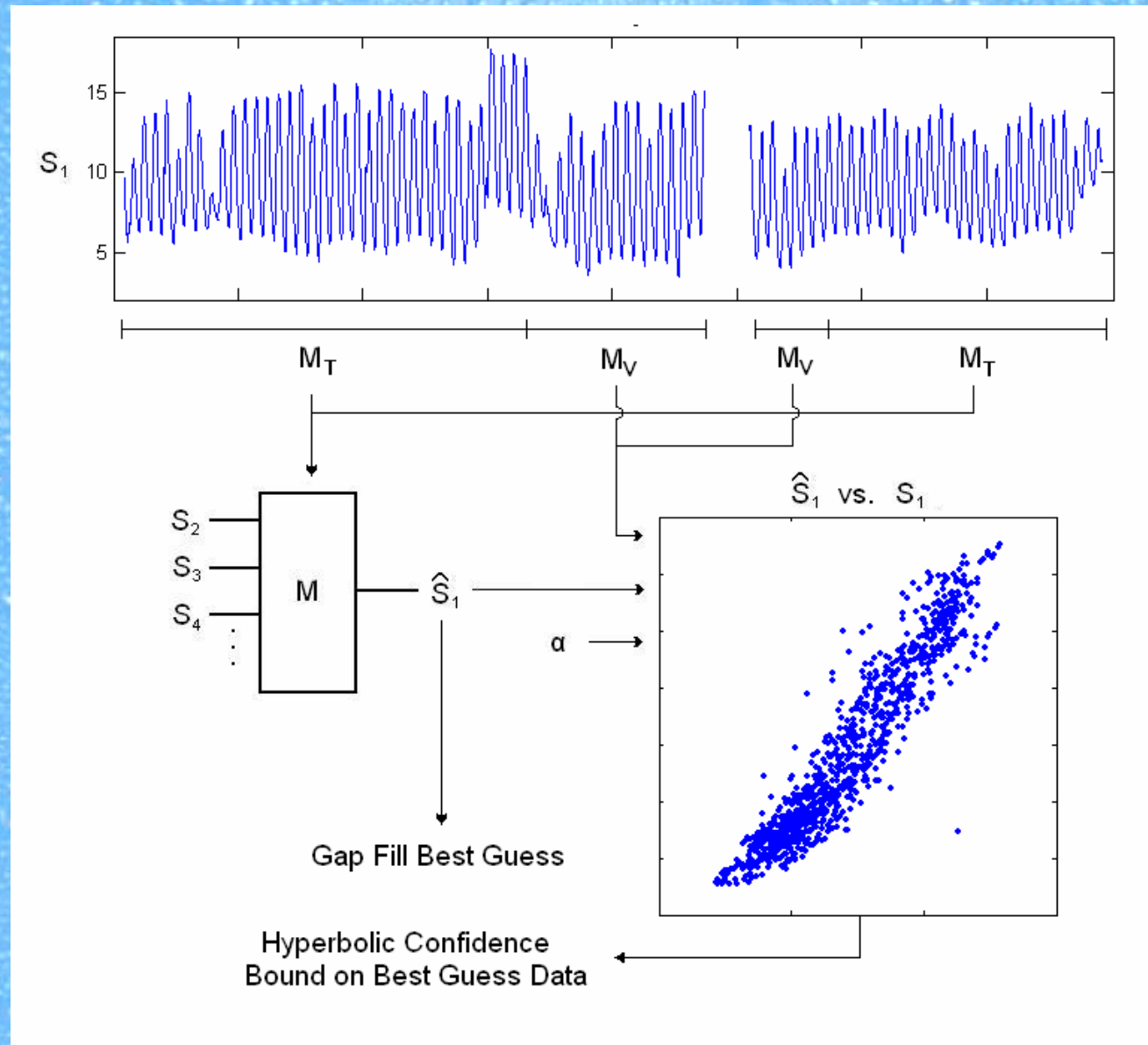
Model Residual Normalization: Residual

- Residual Colouration Correction

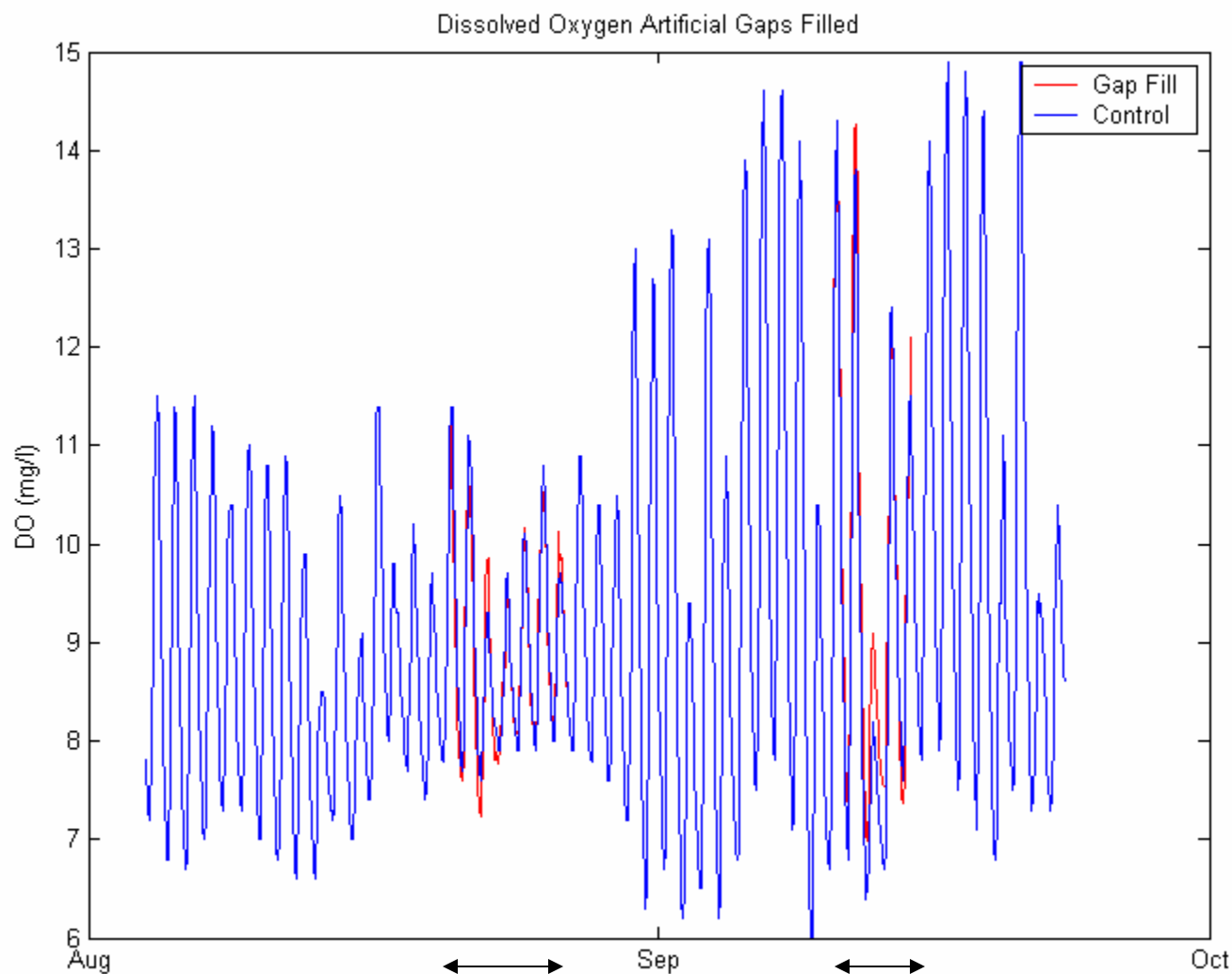
Let's do it...



Data Correction: Gap Filling Methodology

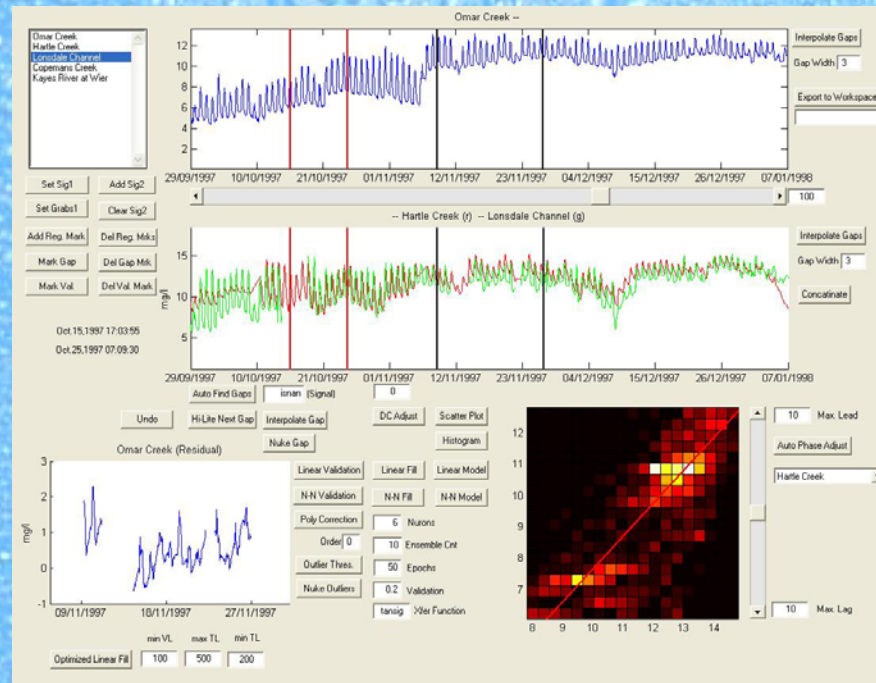


Data Correction: Gap Filling Control

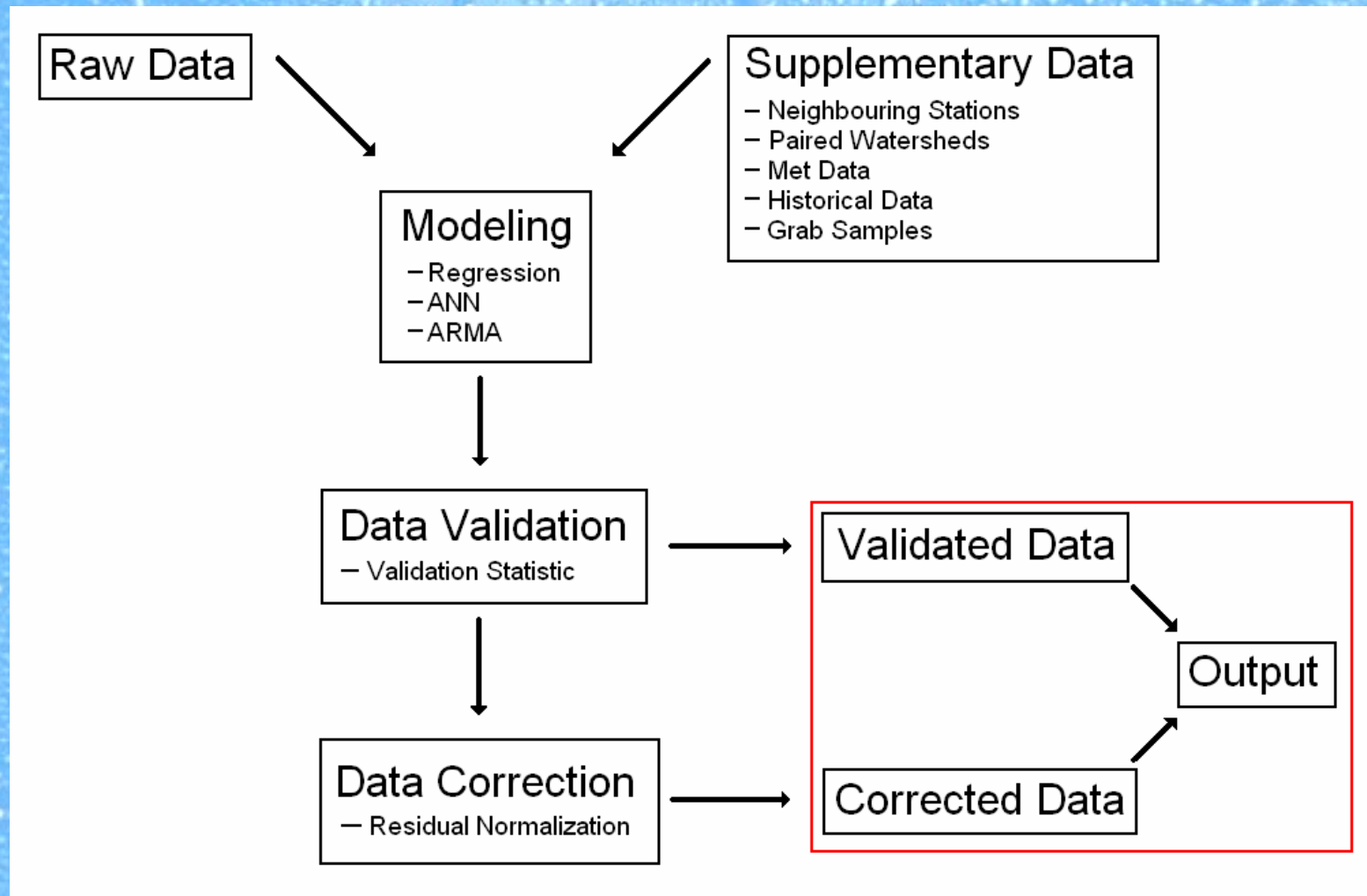


Data Correction: Gap Filling

Let's Do it...



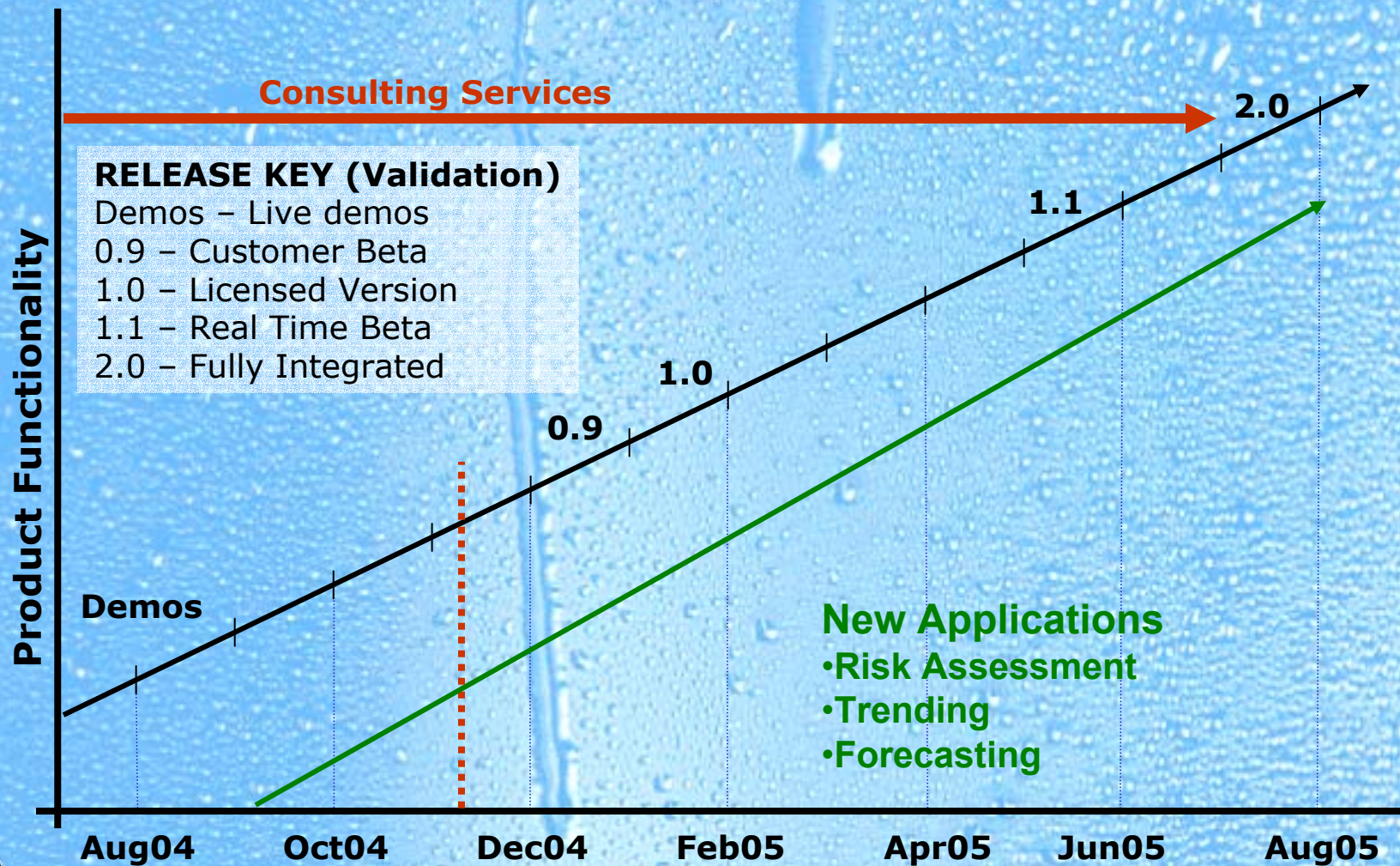
Data Management Flow Chart: Data Output

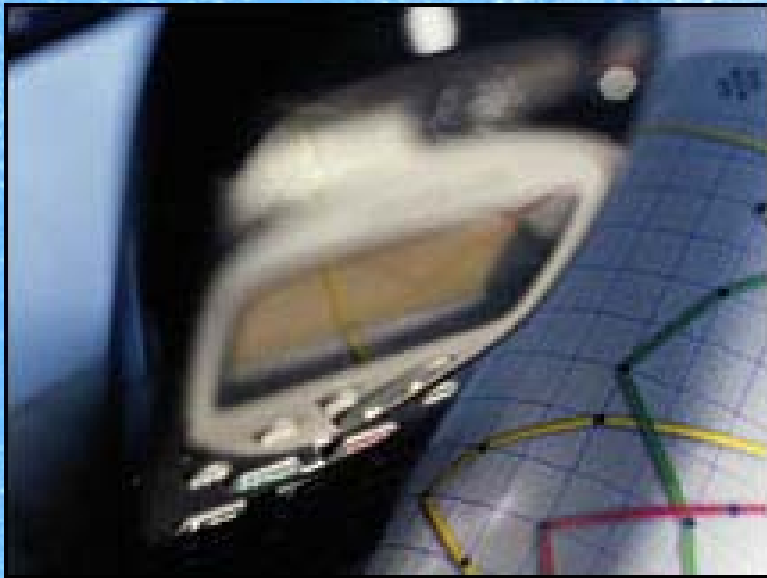


Data Output: Validated / Corrected Data

Untitled - Notepad							
File Edit Format View Help							
Station: Lang Creek at Hatchery							
Parameter: Temperature							
Units: °C							
Start Date & Time: March-15-1999 15:47:13							
End Date & Time: July-01-2000 14:32:13							
Time Stamp Interval: 00:15:00							
Validation & Correction Performed By: Touraj Farahmand							
Validation & Correction Performed On: October-17-2004							
Validation Stat Algorithm Parameters							
nV = 150;							
nT = 300;							
MES = 5;							
MT = Linear;							
Date & Time	Raw Signal,	val Stat,	CType	Correction,	signal out,	Mod 95% Conf,	
March-15-1999 15:47:13	12.3	0.23	0	0	12.3	0	
March-15-1999 16:02:13	12.3	0.23	0	0	12.3	0	
March-15-1999 16:17:13	12.3	0.23	0	0	12.3	0	
March-15-1999 16:32:13	12.4	0.27	0	0	12.4	0	
March-15-1999 16:47:13	12.4	0.27	0	0	12.4	0	
March-15-1999 17:02:13	12.4	0.26	0	0	12.4	0	
March-15-1999 17:17:13	12.3	0.24	0	0	12.3	0	
March-15-1999 17:32:13	12.3	0.24	0	0	12.3	0	
March-15-1999 17:47:13	12.3	0.23	0	0	12.3	0	
March-15-1999 18:02:13	12.3	0.23	0	0	12.3	0	
March-15-1999 18:17:13	27.4	12.99	2	12.33	12.33	1.2	
March-15-1999 18:32:13	12.3	0.23	0	0	12.3	0	
March-15-1999 18:47:13	12.3	0.23	0	0	12.3	0	
March-15-1999 19:02:13	12.2	0.12	0	0	12.3	0	
March-15-1999 19:17:13	12.2	0.12	0	0	12.3	0	
March-15-1999 19:32:13	12.2	0.11	0	0	12.3	0	
March-15-1999 19:47:13	12.2	0.11	0	0	12.3	0	
March-15-1999 20:02:13	12.3	0.23	0	0	12.3	0	
March-15-1999 20:17:13	15.3	5.32	1	-3.1	12.2	1.2	
March-15-1999 20:32:13	15.3	5.32	1	-3.1	12.2	1.2	
March-15-1999 20:47:13	15.3	5.32	1	-3.1	12.3	1.2	
March-15-1999 20:02:13	15.3	5.32	1	-3.1	12.2	1.2	
March-15-1999 21:17:13	15.3	5.32	1	-3.1	12.2	1.2	
March-15-1999 21:32:13	15.3	5.32	1	-3.1	12.2	1.2	
March-15-1999 21:47:13	15.3	5.32	1	-3.1	12.2	1.2	
March-15-1999 21:02:13	15.4	5.32	1	-3.1	12.3	1.2	
March-15-1999 22:17:13	15.3	5.32	1	-3.1	12.2	1.2	
March-15-1999 22:32:13	12.3	0.23	0	0	12.3	0	
March-15-1999 22:47:13	12.3	0.23	0	0	12.3	0	
.							
.							
.							
.							

AI Roadmap





Peter Hudson

pete@aquaticinformatics.com

